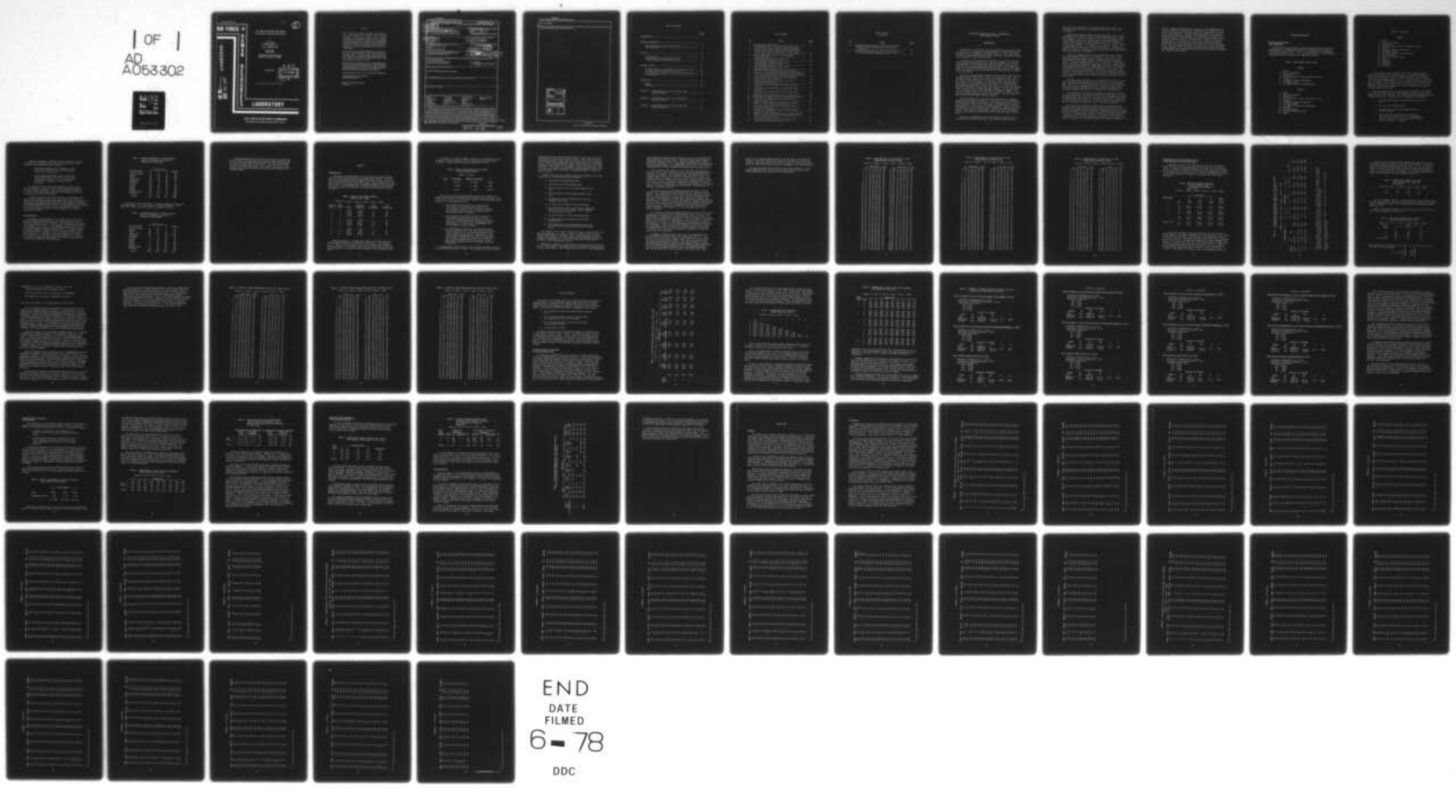
AD-A053 302 APPLIED SCIENCE ASSOCIATES INC VALENCIA PA F/G 5/10
JET ENGINE MECHANIC-AFSC 426X2: EXPERIMENTAL JOB PERFORMANCE TE--ETC(U)
DEC 77 J K HAWLEY, C J MULLINS, J WEEKS F33615-76-C-0057
UNCLASSIFIED AFHRL-TR-77-73 NL

| OF |
AD
A053302



END
DATE
FILMED
6-78
DDC

AFHRL-TR-77-73

AIR FORCE

HUMAN RESOURCES

AD A053302

AD No.
DDC FILE COPY

JET ENGINE MECHANIC—AFSC 426X2: EXPERIMENTAL JOB PERFORMANCE TESTS

By

John K. Hawley
Applied Science Associates, Inc.
Box 158
Valencia, Pennsylvania 16059

Cecil J. Mullins
Joseph Weeks

PERSONNEL RESEARCH DIVISION
Brooks Air Force Base, Texas 78235

December 1977

LABORATORY

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235

NOTICE

This final report was submitted by Applied Science Associates, Inc., Box 158, Valencia, Pennsylvania 16059, under contract F33615-76-C-0057, project 7719, with Personnel Research Division, Air Force Human Resources Laboratory (AFSC), Brooks Air Force Base, Texas 78235. Mr. Joseph Weeks, Manpower Development and Evaluation Branch, was the contract monitor.

This technical report has been reviewed and is approved for publication.

LELAND D. BROKAW, Technical Director
Personnel Research Division

DAN D. FULGHAM, Colonel, USAF
Commander

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (*When Data Entered*)

19 REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 18 AFHRL TR-77-73 ✓	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) 6 JET ENGINE MECHANIC–AFSC 426X2: EXPERIMENTAL JOB PERFORMANCE TESTS.		5. TYPE OF REPORT & PERIOD COVERED 9 Final rept.,
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) 10 John K. Hawley, Cecil J. Mullins Joseph Weeks		8. CONTRACT OR GRANT NUMBER(s) 15 F33615-76-C-0057 new
9. PERFORMING ORGANIZATION NAME AND ADDRESS Applied Science Associates, Inc. Box 158 Valencia, Pennsylvania 16059 ✓		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62703F 17 14 16 77191403
11. CONTROLLING OFFICE NAME AND ADDRESS HQ Air Force Human Resources Laboratory (AFSC) Brooks Air Force Base, Texas 78235		12. REPORT DATE 11 Dece[redacted]77 12 73p. 13. NUMBER OF PAGES 72
14. MONITORING AGENCY NAME & ADDRESS(*if different from Controlling Office*) Personnel Research Division Air Force Human Resources Laboratory Brooks Air Force Base, Texas 78235		15. SECURITY CLASS. (*of this report*) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (*of this Report*)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (*of the abstract entered in Block 20, if different from Report*)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (*Continue on reverse side if necessary and identify by block number*)

ASVAB	item analysis	multiple correlation	technical school grades
coefficient alpha	item difficulty	phi-coefficient	validity
correlation	item variance	reliability	
criterion test	Job Performance Test	skill level	
cross-validation	KR20	stepwise regression	

20. ABSTRACT (*Continue on reverse side if necessary and identify by block number*)

Historically, the Air Force has used technical school grades (TSGs) to validate aptitude tests. The purpose of the current study was to develop a job-related criterion metric against which to validate aptitude measures. Along this line, three Criterion Tests relevant to the 3, 5, and 7 skill level of PAFSC 426X2–jet engine mechanic–were developed. ANOVA results demonstrated that mean Criterion Test scores were significantly different across the relevant skill levels. Stepwise regression results indicated that Armed Services Vocational Aptitude Battery (ASVAB) information subscales were most predictive of Criterion Test performance for experienced mechanics. Considering only basic airmen, the ASVAB information measures and general knowledge subscales were most related to Criterion Test performance. When TSGs were regressed on ASVAB scores, general knowledge scales were most consistently → over

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (*When Data Entered*)

632 170 Gull

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(*When Data Entered*)

Item 20 Continued:

predictive of technical school performance. The regression of TSGs on Criterion Test scores indicated that only Test 1 was generally related to technical school performance.

ACCESSION for	
NTIS	White Section ☑
DDC	Buff Section ☐
UNANNOUNCED	☐
JUSTIFICATION	
BY	
DISTRIBUTION/AVAILABILITY CODES	
Dist.	AVAIL. and/or SPECIAL
A	

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(*When Data Entered*)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES
(continued)

JET ENGINE MECHANIC-AFSC 426X2: EXPERIMENTAL JOB PERFORMANCE TESTS

INTRODUCTION

Historically, the Air Force has used technical school grades (TSGs) as a criterion for validating aptitude measures. However, TSGs have not proven to be a completely acceptable criterion. One of the primary reasons for the unacceptability of TSGs is the weak relationship between school grades and later job performance.

The purpose of this study was to develop a job-related criterion metric against which aptitude tests could be validated. In line with this objective, it was desired that the new criterion metric be more indicative of on-the-job performance than are TSGs. A constraint, imposed by the high developmental and administrative costs of job sample tests, was that the new criterion metric be a paper-and-pencil measure of job performance.

The use of paper-and-pencil measures usually entails lower fidelity in simulating job tasks than do actual job performance items. This decrement usually results from the fact that paper-and-pencil tests do not cover a sufficient number of job-relevant tasks. An approach to the problem of task coverage is to concentrate upon comprehensively identifying job components through task analysis. A second problem with paper-and-pencil items is that they often do not cover the complete range of skills within a particular task category. To insure fidelity, it is necessary that all task categories be adequately sampled, and that the complete range of skills within categories be covered. Each item on the paper-and-pencil measure then maps to a specific task category and skill level within the task category.

Once the relevant job skills have been identified and suitable paper-and-pencil measures developed, a link must still be established between job knowledge, as assessed on the criterion measure, and actual job performance. The current study makes the assumption that career progression is a direct measure of job performance. That is, the selection processes of the Air Force tend to promote job incumbents who successfully meet their job requirements. Similarly, when an incumbent ceases to satisfactorily perform his duties, promotion stops or slows down.

Following the assumption that career progression reflects job performance, the current study makes the further assumption that, to

the extent test performance is associated with career level, the Criterion Tests (CTs) may be considered as indices of actual job proficiency.

Employing an indirect measure of job proficiency such as career progression does, however, entails a certain amount of risk. A primary element of risk is that there may be less relationship between job proficiency and career progression than initially assumed. The factors which account for career progression may only indirectly involve job knowledge. The Criterion Test items may, indeed, reflect job knowledge, but fail to discriminate skill groups. However, if Criterion Tests differentiate various skill groups, they then may be said to at least be a measure of "job success," with job success being defined as long-term job incumbency.

Given the foregoing objectives and assumptions, the Criterion Test development effort proceeds as follows. First, the Air Force Occupational Survey (AFOS) is used to identify the content of Air Force Specialty Code (AFSC) 426X2, Jet Engine Mechanic (formerly 432X0). The task responsibilities of 3, 5, and 7 skill level personnel in this AFSC are separately identified. Finally, a large pool of potential paper-and-pencil test items (70 percent more than planned) are designed in such a way as to discriminate 1 and 3, 3 and 5, and 5 and 7 skill level job incumbents. The ability of the CTs to discriminate career progression levels is the link between test performance and job proficiency. An initial editing and screening of the item pool to remove defective items results in tryout versions of the Criterion Tests.

The next step in the process is to administer the CTs to a large group of subjects representative of each skill level in the target AFSC. Results from this administration are used to select the "best" CTs. Best, in this context, refers to CTs made up of those items having the highest discriminating power, the most homogeneity, and the most acceptable distribution of responses to the distractors.

After defective items have been removed, and the final versions of the CTs assembled, the study proceeds with an investigation of test discriminability and reliability. As previously discussed, the ability of the Criterion Test scores to discriminate career levels links the CT results with job proficiency. It is also desirable to demonstrate that the responses of subjects to the CT items are consistent.

The last phase of the study involves the relationships of the Criterion Tests and the Armed Services Vocational Aptitude Battery (ASVAB), the standard Air Force aptitude measure used in the selection and classification of enlisted personnel. ASVAB-Criterion Test relationships are of interest since selection validity hinges upon being able to predict test performance, and, thus, indirectly, job proficiency, from ASVAB scores. The relationships between TSGs and the ASVAB scales permit a comparison of the present results with previous validity

studies. ASVAB-TSG associations, in comparison with ASVAB-Criterion Test correlations, also give some indication of how changing the criterion for selection will possibly affect the composition of the pool of jet engine mechanics. Finally, should Criterion Test results be associated with career progression, the relationship of TSGs with the Criterion Tests will give an indication of how technical school performance relates to eventual job proficiency. Each of these issues is discussed in further detail in the next two sections. Research Methodology covers the development and evaluation of individual test items. The discriminability and reliability of the tests is addressed in the Results section. The Results section and the Validity Studies section address the relationships among the Criterion Tests, the ASVAB, and technical school grades.

RESEARCH METHODOLOGY

Task Analysis and Item Pool Development

The first step in developing a pool of paper-and-pencil job performance items was to systematically identify the separate task components of AFSC 426X2 using the AFOS. Various job components were then grouped into broad categories of activities called task domains. The task domains identified for each skill level are presented in Table 1.

TABLE 1 TASK DOMAINS OF ASFC 426X2

Level-3

1. Safety
2. Tools and Equipment
3. Inspections
4. Technical Orders, Forms, Maintenance Levels
5. Security
6. Shipping, Storage, Preservation
7. Jet Engine Theory
8. Jet Engine Operation and Maintenance

Level-5

1. Safety
2. Tools and Equipment
3. Inspections
4. Technical Orders, Forms, Maintenance Levels
5. Security
6. Shipping, Storage, Preservation
7. Jet Engine Theory
8. Jet Engine Operation and Maintenance
9. Training
10. Work Performance
11. Non-Work Personnel Issues
12. Supply

TABLE 1 (continued)

Level-7

1. Safety
2. Tools and Equipment
3. Inspections
4. Technical Orders, Forms, Maintenance Levels
5. Security
6. Shipping, Storage, Preservation
7. Jet Engine Theory
8. Jet Engine Operation and Maintenance
9. Work Performance
10. Non-Work Personnel Issues
11. Supply
12. Evaluation
13. Planning
14. Supervision
15. Training

As may be noted, the task domains for each of the three skill levels have some, but not complete, overlap. New domain elements are added as the skill level of job incumbents increases. For example, "Training" is a domain element for level 5, but not level 3. The fact that a particular domain element is on more than one list does not mean, however, that test items representing that element are the same across different skill levels. Clearly, the activities involved in a job element like "Safety" are different for a level 3 mechanic and a level 7 noncommissioned officer (NCO).

Following identification of the relevant task domains, job charts, or systematic matrices of skill components which make up the job of jet engine mechanic, were prepared. The job matrices were compiled separately for each skill level. Specifically, the job matrices identified:

1. The tasks done by most job incumbents at that skill level.

2. The tools and equipment used.

3. The knowledge, skills and judgments needed to perform the task.

4. Tasks generally <u>not</u> performed by mechanics at the next lower skill level, or, if performed at the lower level, done in a way requiring different skills, knowledge, or judgments.

Using the information contained in the job analysis matrices, a large pool of potential test items was then developed. Guidelines for item development were as follows:

1. Items should measure job performance on tasks done by most mechanics at a particular skill level (e.g., 3, 5, and 7).

2. Items should discriminate between the performance of mechanics at one skill level and the performance of those at the next lowest skill level, e.g., 1 versus 3, 3 versus 5, and 5 versus 7.

The unrefined item pool was next submitted to the Air Force for final approval. Test items were then edited and screened to eliminate obvious defectives. The pre-screening process resulted in a pool of 450 items (150 in each test for inclusion in the tryout versions of the Criterion Tests).

It should be noted at this point that the job performance Criterion Tests developed in the current study do not have the same objective as the Specialty Knowledge Tests (SKTs) used in the Weighted Airman Promotion System (WAPS). The SKTs are developed on the basis of training documentation and represent training standards of performance. On the other hand, the Criterion Tests were designed to reflect on-the-job performance--that is, job performance as it actually occurs in the field.

Data Collection

Following Air Force approval of the final sets of test items, and the preparation and printing of the testing materials, data collection began at 12 Air Force bases. The preliminary data collection plan is shown in Table 2. More subjects than actually required were scheduled at each base. Over-scheduling was done to serve as a precaution against collecting unusable data, or failing to obtain the expected number of subjects at some bases. The numbers planned for each base were a function of the actual number of each skill level reported to be available. Testing on each group of subjects was scheduled to be completed during one working day. To control for possible fatigue and order effects, the order of administration of the various tests was counterbalanced.

TABLE 2 PLANNED FREQUENCIES OF SUBJECTS WITHIN PAFSC 426X2 ACROSS SKILL LEVELS AT SELECTED AIR FORCE BASES

Air Force Base	Skill Level 3	5	7	Totals
Seymour Johnson	30	25	35	90
McGuire	25	40	40	105
Dover	35	85	42	162
Little Rock	38	20	32	90
Barksdale	28	18	26	72
Columbus	28	5	28	61
Nellis	25	35	12	72
Travis	80	80	40	200
George	26	9	37	72
Castle	28	70	64	162
Mountain Home	28	28	34	90
Holloman	48	15	45	108
Totals:	419	430	435	1284

As expected, the actual number of subjects obtained was lower than the number sought. The actual numbers of subjects obtained in each skill level at each Air Force base are presented in Table 3.

TABLE 3 OBTAINED FREQUENCIES OF SUBJECTS WITHIN PAFSC 426X2 ACROSS SKILL LEVELS AT SELECTED AIR FORCE BASES

Air Force Base	Skill Level 3	5	7	Totals
Seymour Johnson	0	26	9	35
McGuire	21	20	7	48
Dover	55	72	50	177
Little Rock	29	35	38	102
Barksdale	22	6	21	49
Columbus	9	1	17	27
Nellis	6	31	18	55
Travis	4	46	19	69
George	5	15	15	35
Castle	22	34	30	86
Mountain Home	9	30	29	68
Holloman	12	20	30	62
Totals:	194	336	283	813

During approximately the same time period that job performance data was being collected from active duty jet engine mechanics, the Air Force was also administering each Criterion Test and the ASVAB to 422 basic airmen (level 1 personnel). Results from 419 level 1 subjects were eventually usable. The complete test data file thus contains scores from 1232 subjects--813 experienced mechanics and 419 basic airmen.

RESULTS

Item Selection

Following the administration of the Criterion Tests and the ASVAB, individual results were assembled and prepared for analysis. The first step in the data analysis was to obtain each subject's score on each of the Criterion Tests and the nine ASVAB scales. The Criterion Test scores were corrected for guessing using the standard formula used with multiple-choice tests. The resulting means, standard deviations, 25th, and 75th percentile scores for each group on the three Criterion Tests are given in Table 4.

TABLE 4 SCORES ON THE THREE CRITERION TESTS BY SKILL LEVEL

(Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

Criterion Test	Skill Level	Mean	Standard Deviation	25th Percentile	75th Percentile
1	1	9.41	8.91	3	14
1	3	75.15	21.86	63	90
1	5	72.32	22.25	60	88
1	7	93.57	16.03	86	103
2	1	7.46	8.38	0	12
2	3	21.48	14.04	12	31
2	5	28.13	15.21	18	38
2	7	54.78	17.48	43	66
3	1	3.57	6.94	-2	7
3	3	10.35	7.86	4	15
3	5	16.51	11.69	8	23
3	7	38.00	14.19	27	47

Referring to Table 4, the mean test scores on all three Criterion Tests generally increase as a function of skill level. One exception is CT 1, where level 5 personnel do not score as high on the average as level 3 personnel. The quartile scores for each group indicate that the distributions of test scores are nearly symmetric about individual group means.

An analysis of variance (ANOVA) performed on the three test scores across skill levels indicates that the group means are significantly different. Partial ANOVA results are presented in Table 5.

TABLE 5 PARTIAL ANOVA RESULTS FOR THE THREE 150-ITEM CRITERION TESTS

(Total N = 1,232)

Test	F-Statistic	Degrees of Freedom	P
1	1654.22	3, 1228	<.0001
2	679.83	3, 1228	<.0001
3	632.11	3, 1228	<.0001

The next step in the test development effort was to reduce the number of items in each Criterion Test by selecting the 100 "best" items in each test. "Best," in this context, refers to those items which most closely meet the following criteria:

1. The item discriminates between the appropriate skill levels. For example, good items in CT 1--the criterion for skill level 3--should be answered correctly more often by level 3 personnel than by basic airmen (level 1 personnel).

2. An item should be correlated with the total score. Items having a high correlation with the total score more consistently measure the job proficiency aspects on which the whole test discriminates skill level groups. Job proficiency is considered to be a unidimensional, global construct.

3. The distribution of answers to the item distractors should be satisfactory. This last check will identify items in which one or more response alternatives are rarely selected by the testees. Additionally, tabulating answer distributions will identify items where some incorrect response alternatives are marked more often by high scorers than by low scorers. Items having these characteristics are defective and should be rejected.

In applying the above selection rules, primary importance was given to the discriminating power of items. The rationale for this choice is

based upon the fact that the overall purpose of the tests is to serve as criteria for their associated skill level. Thus, tests that do not discriminate between career levels are not suited for use as criteria. If the items do not discriminate career levels, then the link between the tests and job proficiency is broken. Selection rules 2 and 3 were included to lend support to the primary criterion for selection. These rules were to be used in the event that item discrimination alone did not provide sufficient information for deciding whether to retain or reject an item.

Items on the three Criterion Tests were next submitted to an item analysis and the following statistics obtained:

1. The index of discrimination (phi).
2. The rank of the discrimination index.
3. The item difficulty for the group relevant to the test.
4. The item variance for the group relevant to the test.
5. The item-total score correlation for the group relevant to the test.
6. The rank of the item-total score correlation.
7. The item difficulty for the skill group immediately lower than the criterion group. This is denoted the "non-relevant 'group.'"
8. The item variance for the non-relevant group.
9. The item-total score correlation for the non-relevant group.
10. The average rank of the discrimination index and the item-total score correlation for the criterion group.

The complete lists of item statistics for each of the tests are presented in Appendices 1, 2, and 3. Items are listed in order of their discriminating power, as indicated by phi. Also included in the last column of each appendix is a table entry indicating whether the item was selected for inclusion in the final versions of the Criterion Tests. A "Yes" or a "No" indicates this choice.

Referring to Appendix 1, which contains the item analysis statistics for CT 1, the indices of discrimination (phi coefficients) computed across items for skill levels 1 and 3 range from 0.64 to 0.17. The corresponding

item difficulties indicate that 60-90 percent of level 3 personnel generally answer the items correctly. These difficulty levels indicate a desirable situation, since CT 1 is designed to differentiate level 3 from level 1 personnel. The corresponding item-total score correlations for the retained items on CT 1 range from 0.49 to 0.14. None of the retained items in CT 1 has a negative correlation with the total score.

The item analysis statistics on the second test (see Appendix 2) indicate that CT 2 does a relatively poor job of discriminating level 5 from level 3 personnel. The discrimination coefficients for the two relevant groups (skill levels 3 and 5) range from 0.21 to -0.23. For retained items, the range is from 0.21 to .01. Given the sample size involved (N = 530), a value of at least 0.085 is required for a correlation coefficient to be considered significantly different from zero at the $p=.05$ level of significance. Using the $p=.05$ criterion, only 31 test items on CT 2 yield a discrimination index which is individually significantly different from zero.

The reason for the low discriminating power of items in CT 2 is apparent when considering the item difficulties for the two relevant groups. The problem is that most items on CT 2 are generally too difficult for level 5 personnel. In most cases, level 5 personnel answer the items correctly more often than level 3 personnel. However, the magnitude of the difference is not large enough to reliably differentiate the two groups. There is too much overlap in the response distributions. The level of response overlap is evident from the quartile scores presented in Table 4. Here, it is noted that more than 25 percent of level 3 personnel score higher on CT 2 than the mean score of the level 5 personnel.

CT 3 presents a situation midway between the relatively good characteristics of CT 1 and the poor characteristics of CT 2. On CT 3, the discrimination coefficients for retained items range from 0.48 to .01. Again, given the sample size involved (N = 619), a correlation of approximately 0.08 is required for significance at the $p=.05$ level. Using that criterion, 85 items on CT 3 individually yield significant indices of discrimination. Looking at the corresponding item difficulty statistics for items with $\phi < .08$, it is again the case that items on CT 3 having low discriminating power suffer from the same problem as the majority of items on CT 2: the items are generally too difficult for the criterion skill level personnel.

In summary, Criterion Tests 2 and 3 do a less satisfactory job of differentiating their relevant skill groups than does CT 1. On both CTs 2 and 3, the major problem concerns items that are too difficult for the criterion skill groups. This result may indicate that differentiating within skill levels of jet engine mechanics is much more difficult than simply differentiating mechanics from non-mechanics. One of the difficulties encountered in the item development phase of the study was deciding specifically which task elements separated level 3 jobs from level 5 jobs, and level 5 jobs from level 7 jobs.

There were no clearly defined criteria. In most cases, job content differences across skill levels were a matter of the degree of the skills involved. This overlap in the job content of the 3, 5, and 7 level skill groups is reflected in their respective performance on the Criterion Tests.

To recap the item selection phase of the study, Tables 6, 7, and 8 present sequential lists of all retained Criterion Test items and the corresponding difficulty indices for all skill levels.

TABLE 6 ITEMS RETAINED IN CRITERION TEST 1 WITH CORRESPONDING DIFFICULTY LEVELS

(Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

		Skill Level				Pooled			Skill Level				Pooled
	Item	1	3	5	7	3-5-7		Item	1	3	5	7	3-5-7
1	2	.24	.79	.62	.74	.70	51	81	.26	.93	.94	.98	.95
2	3	.27	.77	.69	.91	.79	52	82	.43	.89	.88	.96	.91
3	6	.15	.45	.43	.77	.55	53	83	.42	.81	.85	.92	.86
4	8	.24	.87	.82	.91	.86	54	87	.27	.83	.88	.97	.90
5	9	.21	.85	.87	.95	.90	55	89	.24	.84	.80	.89	.84
6	10	.19	.81	.84	.89	.85	56	90	.23	.86	.89	.97	.91
7	14	.18	.49	.54	.72	.59	57	91	.44	.93	.93	.99	.95
8	15	.20	.71	.61	.92	.74	58	92	.23	.61	.62	.86	.70
9	17	.27	.86	.90	.87	.88	59	93	.27	.59	.52	.51	.53
10	18	.54	.95	.92	.97	.95	60	94	.30	.65	.70	.89	.76
11	21	.52	.91	.85	.95	.90	61	95	.25	.91	.90	.98	.93
12	23	.14	.64	.66	.88	.73	62	96	.54	.86	.87	.93	.89
13	26	.41	.60	.61	.93	.72	63	97	.29	.65	.70	.87	.75
14	27	.16	.54	.64	.78	.66	64	98	.21	.64	.54	.65	.60
15	28	.40	.88	.91	.87	.89	65	99	.31	.89	.82	.94	.88
16	31	.14	.36	.43	.83	.55	66	101	.19	.52	.40	.44	.44
17	32	.26	.80	.80	.85	.82	67	102	.41	.87	.79	.86	.83
18	34	.33	.91	.90	.98	.93	68	104	.36	.79	.79	.88	.82
19	35	.18	.79	.80	.86	.82	69	105	.22	.41	.48	.80	.57
20	36	.16	.74	.79	.83	.79	70	106	.27	.68	.63	.72	.68
21	37	.18	.66	.70	.92	.77	71	108	.18	.54	.36	.45	.44
22	40	.62	.95	.93	.96	.95	72	110	.19	.47	.31	.52	.42
23	41	.25	.85	.87	.95	.90	73	111	.26	.75	.69	.79	.74
24	42	.30	.81	.84	.93	.87	74	112	.23	.72	.68	.81	.74
25	43	.38	.81	.80	.88	.83	75	114	.41	.60	.73	.87	.75
26	44	.32	.70	.64	.73	.69	76	115	.62	.94	.87	.98	.92
27	46	.36	.75	.75	.88	.80	77	116	.17	.70	.68	.57	.65
28	47	.11	.29	.19	.82	.43	78	117	.34	.77	.84	.89	.84
29	48	.37	.89	.91	.95	.92	79	119	.26	.54	.48	.57	.53
30	49	.36	.71	.68	.78	.72	80	120	.23	.74	.71	.92	.79
31	52	.11	.61	.54	.59	.58	81	121	.21	.53	.33	.28	.36
32	54	.41	.93	.87	.97	.92	82	123	.26	.70	.61	.77	.69
33	55	.24	.72	.75	.84	.78	83	124	.23	.58	.62	.83	.69
34	58	.33	.74	.65	.79	.72	84	126	.21	.81	.80	.90	.84
35	59	.39	.89	.82	.90	.86	85	128	.48	.80	.78	.96	.85
36	61	.24	.66	.77	.94	.80	86	129	.11	.60	.65	.85	.71
37	62	.13	.71	.66	.78	.72	87	131	.17	.63	.39	.52	.49
38	63	.59	.96	.89	.98	.94	88	132	.30	.75	.67	.84	.75
39	64	.30	.81	.83	.95	.87	89	134	.22	.59	.38	.22	.38
40	65	.34	.76	.78	.80	.78	90	135	.25	.47	.43	.49	.46
41	66	.42	.70	.60	.56	.61	91	136	.10	.46	.52	.56	.52
42	68	.35	.76	.78	.89	.81	92	137	.18	.41	.52	.83	.60
43	69	.25	.79	.85	.95	.87	93	138	.27	.56	.62	.70	.63
44	70	.38	.84	.78	.75	.78	94	140	.14	.38	.38	.48	.42
45	71	.29	.88	.87	.95	.90	95	141	.17	.76	.77	.82	.79
46	72	.44	.90	.90	.96	.92	96	142	.15	.66	.58	.64	.62
47	74	.31	.72	.64	.69	.68	97	145	.20	.55	.49	.62	.55
48	75	.20	.71	.62	.65	.65	98	146	.11	.56	.60	.81	.66
49	76	.21	.66	.42	.42	.48	99	149	.24	.48	.50	.59	.53
50	77	.27	.88	.89	.98	.92	100	150	.33	.66	.71	.73	.70

TABLE 7 ITEMS RETAINED IN CRITERION TEST 2 WITH CORRESPONDING DIFFICULTY LEVELS

(Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

		Skill Level				Pooled
	Item	1	3	5	7	3-5-7
1	2	.13	.30	.36	.48	.39
2	3	.11	.19	.31	.44	.32
3	4	.23	.28	.31	.30	.30
4	5	.25	.39	.46	.66	.51
5	7	.17	.08	.12	.12	.11
6	9	.29	.53	.67	.92	.72
7	12	.35	.43	.51	.58	.51
8	13	.39	.39	.44	.73	.53
9	14	.27	.22	.29	.72	.42
10	16	.20	.12	.19	.65	.33
11	18	.16	.30	.37	.66	.46
12	19	.43	.57	.59	.82	.67
13	20	.29	.18	.21	.40	.27
14	22	.36	.55	.67	.89	.72
15	24	.30	.27	.37	.67	.45
16	25	.38	.55	.64	.76	.66
17	26	.19	.27	.29	.49	.36
18	27	.12	.11	.21	.39	.25
19	29	.32	.56	.59	.59	.58
20	30	.44	.61	.62	.73	.66
21	31	.18	.61	.67	.91	.74
22	32	.23	.06	.08	.08	.08
23	33	.45	.65	.72	.92	.77
24	34	.22	.30	.38	.62	.44
25	35	.18	.12	.14	.14	.14
26	37	.58	.76	.81	.92	.84
27	38	.15	.45	.50	.67	.55
28	39	.47	.52	.62	.82	.66
29	42	.31	.29	.33	.50	.38
30	43	.38	.63	.74	.88	.76
31	45	.19	.09	.13	.27	.17
32	46	.11	.19	.23	.63	.36
33	47	.29	.20	.22	.43	.29
34	48	.29	.36	.43	.66	.49
35	50	.44	.75	.82	.95	.85
36	51	.29	.46	.55	.71	.58
37	52	.20	.19	.22	.34	.26
38	55	.20	.38	.43	.56	.46
39	57	.33	.25	.34	.73	.46
40	59	.49	.60	.63	.71	.65
41	60	.42	.48	.61	.70	.61
42	62	.25	.49	.63	.90	.69
43	64	.26	.30	.40	.53	.42
44	65	.27	.25	.44	.42	.39
45	66	.20	.15	.18	.33	.23
46	67	.52	.69	.68	.89	.76
47	68	.24	.16	.22	.18	.19
48	69	.32	.42	.46	.79	.56
49	71	.23	.24	.29	.48	.35
50	74	.26	.40	.44	.24	.36
51	75	.19	.21	.24	.34	.27
52	76	.43	.80	.81	.95	.86
53	77	.31	.34	.36	.49	.40
54	78	.24	.30	.43	.73	.50
55	80	.36	.57	.76	.94	.78
56	82	.57	.66	.71	.89	.76
57	83	.52	.59	.67	.70	.66
58	84	.15	.18	.22	.36	.26
59	85	.28	.17	.20	.41	.27
60	86	.18	.58	.60	.72	.64
61	88	.47	.55	.59	.74	.63
62	89	.41	.60	.63	.78	.67
63	90	.32	.55	.56	.88	.67
64	92	.29	.27	.37	.27	.31
65	93	.30	.30	.35	.70	.46
66	99	.28	.40	.54	.80	.60
67	100	.26	.35	.36	.66	.46
68	101	.16	.17	.18	.36	.24
69	103	.26	.35	.42	.62	.47
70	104	.27	.27	.40	.80	.51
71	105	.37	.38	.42	.59	.47
72	106	.26	.30	.35	.37	.34
73	107	.25	.26	.33	.26	.29
74	108	.25	.32	.35	.45	.38
75	109	.33	.38	.43	.62	.48
76	111	.30	.26	.30	.44	.34
77	112	.38	.45	.52	.69	.56
78	114	.41	.64	.64	.71	.66
79	116	.31	.49	.60	.70	.61
80	117	.34	.32	.35	.49	.39
81	118	.24	.43	.45	.52	.47
82	119	.23	.20	.21	.16	.19
83	121	.38	.31	.40	.78	.51
84	123	.17	.25	.41	.81	.51
85	124	.33	.55	.65	.75	.66
86	127	.43	.31	.36	.32	.34
87	129	.28	.37	.49	.58	.49
88	130	.17	.24	.33	.57	.39
89	131	.27	.20	.34	.59	.40
90	133	.20	.17	.22	.26	.22
91	134	.27	.59	.66	.76	.68
92	135	.39	.51	.56	.93	.68
93	136	.14	.19	.19	.42	.27
94	138	.15	.29	.48	.75	.53
95	139	.38	.28	.32	.34	.32
96	142	.41	.41	.48	.52	.48
97	144	.15	.17	.21	.24	.21
98	145	.22	.23	.35	.62	.42
99	147	.12	.24	.36	.40	.34
100	149	.43	.36	.42	.73	.51

TABLE 8 ITEMS RETAINED IN CRITERION TEST 3 WITH CORRESPONDING DIFFICULTY LEVELS

(Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

		Skill Level				Pooled
	Item	1	3	5	7	3-5-7
1	2	.27	.40	.57	.86	.63
2	4	.35	.38	.34	.49	.40
3	6	.29	.51	.49	.60	.53
4	7	.28	.31	.36	.36	.35
5	8	.34	.32	.39	.59	.44
6	9	.21	.22	.22	.28	.24
7	10	.16	.15	.08	.20	.14
8	11	.31	.20	.26	.47	.32
9	12	.18	.22	.15	.27	.21
10	13	.10	.09	.11	.24	.15
11	14	.28	.33	.43	.65	.48
12	15	.15	.16	.13	.17	.15
13	16	.25	.30	.23	.33	.28
14	17	.23	.26	.34	.70	.44
15	18	.24	.20	.24	.70	.39
16	19	.13	.18	.18	.48	.28
17	20	.25	.52	.73	.94	.75
18	21	.24	.27	.37	.59	.42
19	22	.38	.27	.24	.46	.32
20	24	.30	.44	.58	.65	.57
21	25	.22	.24	.20	.66	.37
22	27	.67	.79	.86	.97	.88
23	28	.30	.48	.48	.61	.52
24	29	.37	.53	.71	.89	.73
25	31	.17	.19	.16	.44	.26
26	32	.16	.32	.47	.54	.46
27	33	.39	.48	.48	.48	.48
28	34	.23	.25	.27	.36	.30
29	35	.24	.26	.24	.65	.39
30	39	.31	.34	.42	.72	.51
31	40	.16	.11	.12	.41	.22
32	41	.22	.34	.34	.51	.40
33	43	.23	.38	.45	.72	.52
34	44	.27	.20	.36	.67	.43
35	45	.16	.45	.60	.93	.68
36	46	.18	.15	.09	.13	.12
37	47	.25	.15	.14	.45	.25
38	49	.19	.17	.24	.35	.26
39	51	.50	.57	.67	.81	.69
40	52	.20	.32	.29	.39	.33
41	53	.23	.15	.17	.38	.24
42	54	.15	.15	.12	.20	.16
43	55	.58	.66	.80	.89	.80
44	56	.39	.53	.57	.71	.61
45	57	.21	.12	.12	.52	.26
46	59	.31	.43	.54	.66	.55
47	62	.28	.39	.52	.78	.58
48	63	.26	.41	.37	.58	.46
49	64	.29	.38	.52	.68	.54
50	66	.42	.53	.63	.76	.65
51	68	.25	.61	.60	.80	.67
52	69	.26	.23	.28	.42	.31
53	72	.21	.19	.21	.45	.29
54	73	.32	.38	.50	.71	.55
55	75	.32	.29	.32	.52	.38
56	77	.29	.15	.15	.24	.18
57	79	.35	.41	.39	.48	.43
58	80	.20	.24	.30	.58	.38
59	81	.26	.43	.63	.89	.67
60	82	.26	.34	.37	.43	.38
61	83	.28	.23	.32	.49	.36
62	85	.13	.23	.18	.25	.22
63	86	.23	.27	.31	.46	.35
64	87	.24	.21	.36	.59	.40
65	88	.16	.11	.12	.16	.13
66	90	.17	.18	.28	.66	.39
67	91	.27	.38	.45	.67	.51
68	92	.16	.24	.21	.34	.26
69	93	.29	.20	.31	.69	.41
70	101	.25	.19	.26	.72	.40
71	102	.53	.43	.49	.57	.50
72	103	.20	.19	.27	.55	.35
73	104	.23	.25	.39	.55	.41
74	105	.31	.40	.49	.49	.47
75	107	.17	.25	.22	.30	.26
76	108	.41	.46	.57	.65	.57
77	109	.18	.24	.25	.40	.30
78	113	.30	.28	.43	.71	.49
79	115	.27	.26	.26	.50	.34
80	117	.16	.14	.17	.29	.21
81	119	.28	.32	.34	.55	.41
82	120	.25	.29	.27	.29	.28
83	121	.42	.31	.33	.45	.37
84	122	.19	.24	.39	.78	.49
85	123	.45	.50	.64	.82	.67
86	126	.32	.36	.55	.77	.58
87	127	.27	.23	.26	.41	.31
88	129	.15	.34	.43	.64	.48
89	131	.33	.30	.31	.49	.37
90	132	.27	.34	.36	.40	.37
91	135	.32	.35	.42	.47	.42
92	136	.66	.76	.81	.92	.84
93	137	.13	.33	.23	.30	.28
94	138	.26	.23	.34	.43	.35
95	140	.29	.27	.26	.35	.29
96	143	.52	.60	.66	.75	.68
97	144	.20	.12	.12	.27	.17
98	145	.16	.18	.17	.30	.22
99	146	.38	.29	.40	.60	.44
100	149	.31	.52	.54	.64	.57

Discriminability and Reliability of the Final Versions of the Criterion Tests

Following selection of the 100 best items in each of the three Criterion Tests, the tests were again scored using only the retained items. The means and standard deviations for each test, plus the total of the test scores, classified by skill level, are presented in Table 9. The test mean scores shown in Table 9 now have a maximum possible value of 100 rather than 150. The new test scores are also corrected for guessing.

TABLE 9 MEANS AND STANDARD DEVIATIONS OF TOTAL SCORES ON TESTS 1, 2, 3, AND POOLED TEST SCORE BY SKILL LEVEL
(Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

		Test Score			
Skill Level		1	2	3	Total
1	M	3.98	5.34	3.13	12.45
	SD	7.48	6.56	5.78	13.40
3	M	61.95	15.35	8.54	85.84
	SD	17.89	11.70	7.18	29.56
5	M	59.06	24.00	17.78	100.84
	SD	17.84	13.10	11.10	35.23
7	M	73.46	45.62	38.57	157.65
	SD	12.27	13.65	14.24	34.07
Pooled 3-5-7	M	64.76	29.46	21.57	115.79
	SD	17.36	17.86	17.15	45.62

As was the case before final item selection, the total scores in Table 9 generally increase as a function of increasing skill level. This result holds over all three Criterion Tests. One exception is, again, CT 1 where level 3 personnel score slightly higher than level 5 personnel (61.95 versus 59.06). A multivariate analysis of variance (MANOVA) performed on the three individual test scores indicates that the group mean profiles are significantly different across skill levels ($p < .0001$). A corresponding univariate analysis of variance on the sum of the test scores was also highly significant ($F=1547.71$, with 3 and 1228 DF, $p < .0001$). The analysis of variance results are summarized in Table 10.

TABLE 10. UNIVARIATE AND MULTIVARIATE ANALYSIS OF VARIANCE RESULTS ON THE THREE CRITERION TEST SCORES AND THEIR SUM

(Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

	Multivariate			Univariate							
				T_1		T_2		T_3		T_t	
Effect	F^a	df	p	F	p	F	p	F	p	F	p
Main[b]	3018.67	9	.0001	1844.29	.0001	750.86	.0001	749.86	.0001	1547.71	.0001
3 vs 1	877.42	3,1226	.0001	2336.35	.0001	105.54	.0001	38.96	.0001	888.78	.0001
5 vs 3	48.08	3,1226	.0001	5.41	.019	73.18	.0001	48.05	.0001	22.02	.0001
7 vs 5	316.90	3,1226	.0001	167.14	.0001	570.66	.0001	871.96	.0001	683.89	.0001

[a]The test statistic for the multivariate test of the main effect is an asymptotic chi-square developed from a likelihood-ratio criterion. The multivariate group contrasts are tested with an exact F developed from the same likelihood-ratio criterion.

[b]The tests of the main effect have 3 degrees of freedom, and the error terms have 1228 degrees of freedom in all cases.

The analysis of variance results indicate that the test scores do discriminate between personnel at each of the skill levels. Correlation coefficients computed between skill levels (e.g., 1, 3, 5, and 7) and scores on each of the three Criterion Tests and their sum are shown in Table 11. The correlations lend support to the analysis of variance results, and also give an idea of the degree of association between individual test performance and corresponding skill level.

TABLE 11 CORRELATIONS BETWEEN TEST SCORES AND ASSOCIATED SKILL LEVEL FOR ALL SUBJECTS

(Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

T_1	T_2	T_3	T_T
.851	.770	.723	.876

One of the primary objectives of the Criterion Test development effort has thus been achieved: the test results discriminate very well among skill levels. Individual Criterion Test Performance is also highly associated with skill level.

Table 12 presents reliability coefficients for each of the three Criterion Tests categorized by skill level.

TABLE 12 RELIABILITY COEFFICIENTS ON TESTS 1, 2, AND 3 CLASSIFIED BY SKILL LEVEL

(Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

Skill Level	Test 1	Test 2	Test 3
1	.67	.55	.44
3	.95	.86	.63
5	.95	.88	.85
7	.92	.91	.91
Pooled 3-5-7	.95	.94	.94

The reliability coefficients shown in Table 12 are the KR20 version of Cronbach's coefficient Alpha:

$$(1) \quad r_{tt} = \frac{I}{I-1}\left[1 - \frac{\sum_{i=1}^{I} \sigma_i^2}{\sigma_x^2}\right]$$

In Equation 1, I is the total number of items in the test,

σ_i^2 is the variance of each individual item i,

$\sigma_i^2 = P_i(1-P_i)$ where P_i is the proportion of respondents who answer item i correctly = difficulty of item i,

and σ_x^2 is the variance of the distribution of total scores.

Due to time and scheduling constraints, the current situation permitted only one administration of each of the Criterion Tests. One test administration permits reliability assessment using either a split-half procedure or coefficient alpha (Kuder-Richardson reliability). Split-half reliability estimation was ruled out because of the difficulties involved in splitting the Criterion Tests into equivalent or "parallel" halves. In order to construct parallel halves, each test would have had to be divided into two parts, with each part being equivalent in terms of content, item difficulty, and so forth. In all likelihood, this could have been accomplished; however, a significant increase in development costs would have resulted. Additional costs would also have resulted from the practical necessity that the parallelism of the tests be empirically demonstrated.

Coefficient alpha, or KR20, provides information about the consistency of individual performance from item to item within a single test. The statistic thus provides an estimate of the internal equivalence of the tests. Coefficient alpha has been shown to be a lower bound estimate of the true reliability coefficient of a test. That is, coefficient alpha will always be lower than the true reliability coefficient, and lower than a reliability estimate obtained using another procedure such as split-half or parallel forms. Thus, should coefficient alpha provide an "acceptable" reliability estimate, it may then be said that the true reliability of the test is also "acceptable."

Given the assumption that job proficiency is a unidimensional construct, assessing intra-subject response consistency using coefficient alpha is acceptable. However, if job proficiency is, in reality, a complex or multidimensional construct, then stressing internal consistency using coefficient alpha may not be appropriate. The scope of the present study does not include assessing the factorial nature of the Criterion Tests, or investigating the underlying structure of the concept of job proficiency. The KR20 reliability coefficients are included only as descriptive indices of the degree of intra-subject response consistency across the various items in the tests.

Considering the reliability coefficients presented in Table 12, several patterns are apparent which are worth noting. First, each of the tests provides high reliability/response consistency estimates for its criterion group (e.g., T_1 for level 3 personnel, etc.). Second, as expected, response consistency decreases as the overall difficulty level of the tests increases.

As a final word on reliability/consistency, and also to summarize the descriptive phase of the study, Tables 13, 14, and 15 present item-total score correlations for each of the items on the three Criterion Tests. The correlations are also presented by skill level. The results shown in Tables 13, 14, and 15 extend the information supplied by the KR20 reliability/consistency estimates. That is, CT 1, which yields the highest reliability estimates, also has the highest and most consistent item-total score correlations. This is particularly true for level 3 and level 5 personnel. Item-total score correlations on CTs 2 and 3 are generally lower and less consistent than those of CT 1. This latter situation is also reflected in the lower reliability estimates for CTs 2 and 3.

TABLE 13 ITEM-TOTAL SCORE CORRELATIONS FOR FINAL 100 ITEMS ON TEST 1

(Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

		Skill Level				Pooled			Skill Level				Pooled
	Item	1	3	5	7	3-5-7		Item	1	3	5	7	3-5-7
1	2	.09	.43	.22	.17	.26	51	81	-.00	.20	.19	.18	.21
2	3	.09	.25	.24	.24	.30	52	82	.18	.22	.41	.35	.36
3	6	.05	.20	.31	.23	.35	53	83	.21	.23	.32	.31	.31
4	8	.05	.38	.42	.35	.40	54	87	.11	.30	.32	.23	.33
5	9	.10	.18	.37	.18	.30	55	89	.10	.42	.34	.26	.35
6	10	-.03	.27	.19	.14	.21	56	90	-.06	.38	.38	.14	.37
7	14	.16	.25	.20	.26	.27	57	91	.12	.26	.40	.04	.33
8	15	.20	.27	.24	.27	.34	58	92	.04	.37	.33	.36	.40
9	17	.20	.25	.27	.19	.21	59	93	.04	.22	.19	.11	.15
10	18	.19	.16	.25	.15	.23	60	94	-.01	.33	.35	.28	.38
11	21	.19	.30	.35	.16	.33	61	95	-.07	.48	.41	.01	.39
12	23	.14	.37	.41	.30	.43	62	96	.23	.49	.35	.30	.39
13	26	.08	.39	.28	.26	.39	63	97	.18	.38	.30	.30	.36
14	27	.08	.14	.30	.31	.30	64	98	.13	.28	.33	.40	.34
15	28	.05	.25	.28	.15	.19	65	99	-.05	.43	.32	.37	.39
16	31	.07	.29	.20	.30	.36	66	101	.13	.28	.21	.21	.21
17	32	-.03	.22	.20	-.02	.16	67	102	.05	.19	.26	.21	.23
18	34	.20	.39	.41	.10	.37	68	104	.28	.44	.40	.27	.39
19	35	.08	.17	.35	.15	.26	69	105	.06	.36	.40	.40	.46
20	36	-.03	.19	.07	.26	.16	70	106	.09	.32	.36	.31	.34
21	37	.10	.37	.30	.34	.39	71	108	.13	.21	.18	.18	.19
22	40	.13	.18	.40	.23	.30	72	110	.10	.36	.34	.24	.34
23	41	.05	.32	.39	.34	.38	73	111	.17	.47	.40	.36	.41
24	42	.07	.21	.37	.21	.31	74	112	.19	.39	.47	.33	.43
25	43	.14	.32	.33	.25	.32	75	114	.28	.44	.43	.37	.45
26	44	.17	.29	.21	.20	.24	76	115	.01	.23	.34	.16	.32
27	46	.03	.26	.23	.14	.26	77	116	.22	.35	.32	.16	.22
28	47	.08	.25	.08	.32	.36	78	117	-.21	.23	.36	.13	.28
29	48	.09	.12	.30	.16	.22	79	119	.27	.46	.36	.38	.39
30	49	.17	.29	.23	-.05	.20	80	120	.24	.42	.47	.48	.50
31	52	.11	.34	.19	.11	.21	81	121	.21	.38	.17	.12	.16
32	54	.16	.16	.29	.35	.29	82	123	.25	.38	.42	.51	.45
33	55	.08	.18	.34	.29	.30	83	124	.22	.47	.49	.48	.52
34	58	.15	.27	.31	.17	.29	84	126	.15	.39	.45	.52	.46
35	59	.12	.38	.39	.27	.36	85	128	-.13	.41	.37	.21	.40
36	61	.16	.28	.29	.17	.32	86	129	.04	.42	.45	.50	.49
37	62	.14	.25	.30	.19	.28	87	131	.10	.41	.28	.22	.29
38	63	.14	.21	.23	.29	.26	88	132	.32	.53	.45	.46	.50
39	64	.10	.42	.38	.32	.41	89	134	.22	.26	.25	.19	.14
40	65	.25	.38	.31	.26	.30	90	135	.42	.29	.30	.35	.30
41	66	.20	.22	.22	.10	.15	91	136	.20	.43	.41	.37	.38
42	68	.16	.40	.32	.26	.35	92	137	.22	.32	.37	.41	.44
43	69	.16	.43	.25	.13	.32	93	138	.38	.34	.44	.43	.41
44	70	.13	.43	.24	.15	.22	94	140	.25	.26	.32	.43	.34
45	71	.12	.24	.33	.13	.29	95	141	.24	.34	.46	.52	.43
46	72	.14	.44	.40	.32	.41	96	142	.24	.39	.44	.37	.39
47	74	-.01	.46	.25	.28	.30	97	145	.30	.34	.29	.50	.37
48	75	.08	.35	.29	.13	.24	98	146	.15	.29	.30	.47	.38
49	76	.03	.19	.19	.18	.15	99	149	.24	.30	.29	.20	.28
50	77	.13	.26	.35	.31	.35	100	150	.38	.41	.43	.43	.40

TABLE 14 ITEM-TOTAL SCORE CORRELATIONS FOR FINAL 100 ITEMS ON TEST 2

(Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

		Skill Level				Pooled			Skill Level				Pooled
	Item	1	3	5	7	3-5-7		Item	1	3	5	7	3-5-7
1	2	.03	.28	.18	.24	.26	51	75	.05	-.10	.14	.13	.15
2	3	.05	.04	.18	.23	.26	52	76	.20	.35	.42	.44	.41
3	4	.04	-.06	.13	.03	.05	53	77	.07	.11	.07	.09	.15
4	5	.12	.14	.28	.19	.31	54	78	.08	.16	.30	.50	.47
5	7	.18	.02	-.03	-.13	-.02	55	80	.12	.26	.34	.44	.44
6	9	.08	.28	.33	.34	.44	56	82	.20	.29	.29	.41	.38
7	12	.11	.20	.21	.18	.22	57	83	.17	.29	.28	.18	.23
8	13	.13	.25	.12	.18	.32	58	84	.06	.10	.16	.14	.22
9	14	.06	.27	.20	.26	.46	59	85	.07	.16	.04	.31	.28
10	16	-.00	-.00	.16	.35	.47	60	86	.04	.38	.26	.20	.28
11	18	.11	.22	.22	.31	.39	61	88	.24	.33	.29	.18	.30
12	19	.20	.24	.31	.31	.37	62	89	.18	.41	.36	.30	.36
13	20	.14	.16	.16	.26	.29	63	90	.24	.34	.38	.48	.49
14	22	.17	.32	.33	.30	.42	64	92	.11	.07	.04	.00	-.01
15	24	.03	.37	.19	.22	.39	65	93	.06	.31	.22	.31	.42
16	25	.23	.29	.36	.26	.34	66	99	.12	.23	.34	.49	.46
17	26	.11	.31	.07	.25	.28	67	100	.19	.17	.18	.30	.35
18	27	.01	.10	.13	.27	.30	68	101	.02	-.12	.13	.29	.25
19	29	.12	.13	.21	.27	.16	69	103	.06	.07	.23	.33	.32
20	30	.14	.27	.17	.26	.24	70	104	.11	.22	.36	.53	.55
21	31	.06	.26	.23	.22	.36	71	105	.10	.11	.09	.04	.18
22	32	.08	-.15	-.01	-.01	-.01	72	106	.15	.22	.15	.04	.12
23	33	.22	.32	.37	.38	.42	73	107	-.02	.21	-.01	.16	.05
24	34	.06	.05	.14	.28	.31	74	108	.08	.14	.12	.18	.18
25	35	.14	.07	-.04	.16	.05	75	109	.13	.29	.26	.37	.36
26	37	.20	.32	.36	.27	.34	76	111	.14	.30	.19	.27	.28
27	38	.08	.30	.30	.31	.34	77	112	.20	.32	.28	.28	.34
28	39	.16	.32	.22	.28	.36	78	114	.12	.20	.27	.10	.19
29	42	.08	.19	.24	.19	.27	79	116	.02	.22	.37	.44	.36
30	43	.19	.43	.39	.28	.41	80	117	.15	.12	.16	.29	.25
31	45	-.02	-.01	.09	.20	.23	81	118	.09	.19	.25	.31	.24
32	46	-.05	-.03	.16	.25	.38	82	119	.19	.07	.02	.11	.00
33	47	.04	.05	.14	.15	.24	83	121	.09	.13	.14	.15	.37
34	48	.14	.13	.15	.27	.30	84	123	.03	.24	.35	.52	.56
35	50	.35	.27	.39	.29	.38	85	124	.11	.17	.21	.38	.29
36	51	.15	.34	.20	.11	.28	86	127	.11	.09	.04	-.17	-.03
37	52	.11	.15	.08	.15	.19	87	129	.08	.21	.22	.30	.28
38	55	-.04	.13	.18	.27	.25	88	130	.08	.29	.45	.45	.48
39	57	.10	.07	.12	.27	.39	89	131	.14	.02	.27	.42	.41
40	59	.23	.36	.29	.28	.28	90	133	.11	.02	.01	.29	.14
41	60	.12	.24	.35	.28	.32	91	134	.23	.36	.38	.46	.38
42	62	.08	.35	.31	.24	.44	92	135	.04	.26	.25	.22	.42
43	64	.10	.17	.20	.28	.28	93	136	.03	.15	.22	.40	.37
44	65	.02	.25	.17	.10	.17	94	138	.03	.28	.37	.44	.49
45	66	-.09	.03	-.05	.10	.14	95	139	.14	.06	.01	-.17	-.00
46	67	.18	.31	.28	.25	.34	96	142	.20	.15	.16	-.06	.11
47	68	.05	-.02	.09	-.01	.02	97	144	.02	-.02	.11	.21	.13
48	69	.23	.30	.37	.39	.47	98	145	.05	.17	.22	.28	.38
49	71	.14	-.12	.19	.19	.24	99	147	.06	.14	.17	.24	.21
50	74	.09	.29	.19	-.03	-.02	100	149	.07	.18	.21	.12	.34

TABLE 15 ITEM-TOTAL SCORE CORRELATIONS FOR FINAL 100 ITEMS ON TEST 3

(Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

		Skill Level				Pooled
	Item	1	3	5	7	3-5-7
1	2	.15	.20	.27	.29	.42
2	4	.17	.16	.17	.12	.19
3	6	.22	.19	.22	.31	.24
4	7	-.01	.11	.18	.14	.12
5	8	.16	.09	.16	.30	.30
6	9	.09	-.04	-.01	.14	.09
7	10	.05	-.02	-.03	.03	.09
8	11	.11	.16	.26	.22	.33
9	12	.10	.03	.10	.15	.15
10	13	.04	-.17	.10	.16	.20
11	14	.26	.29	.32	.35	.39
12	15	.02	-.02	-.06	.14	.05
13	16	.08	.20	.07	-.02	.08
14	17	.11	.07	.29	.27	.43
15	18	.12	.16	.37	.23	.50
16	19	.03	-.03	.18	.17	.33
17	20	.11	.10	.26	.25	.38
18	21	.11	.03	.23	.35	.36
19	22	.11	.15	.21	.38	.33
20	24	.13	.21	.10	.25	.22
21	25	.08	-.02	.24	.33	.46
22	27	.18	.34	.31	.21	.32
23	28	.20	.21	.16	.17	.20
24	29	.05	.20	.35	.30	.39
25	31	.07	.01	.16	.30	.34
26	32	.12	.07	.15	.12	.18
27	33	.17	.19	.13	.14	.10
28	34	.11	.12	.14	.11	.16
29	35	.06	.08	.31	.34	.46
30	39	.10	.09	.21	.32	.39
31	40	.06	-.05	.20	.33	.39
32	41	.18	.07	.23	.31	.28
33	43	.10	.13	.20	.15	.32
34	44	.14	.18	.36	.46	.51
35	45	.05	.21	.31	.27	.45
36	46	.10	.02	.00	.12	.05
37	47	.05	-.06	.10	.31	.36
38	49	.05	-.01	.08	.19	.19
39	51	.21	.14	.13	.25	.25
40	52	-.04	-.02	.09	.05	.10
41	53	.01	-.11	.11	.33	.31
42	54	-.04	.10	.16	.04	.13
43	55	.17	.34	.35	.37	.36
44	56	.13	.18	.20	.23	.25
45	57	.01	-.00	.24	.53	.54
46	59	.06	.11	.17	.15	.22
47	62	.14	.18	.28	.31	.40
48	63	.15	.15	.19	.25	.26
49	64	.08	.14	.22	.46	.36
50	66	.15	.30	.40	.28	.35
51	68	.15	.39	.31	.27	.33
52	69	.10	-.06	.20	.19	.23
53	72	-.01	.05	.25	.28	.35
54	73	.12	.08	.20	.35	.35
55	75	.15	.21	.24	.19	.29
56	77	.07	.07	.09	.05	.13
57	79	.12	.25	.31	.23	.23
58	80	.17	.07	.20	.41	.39
59	81	.09	.17	.25	.28	.41
60	82	.10	.22	.24	.19	.19
61	83	.12	-.07	.20	.23	.27
62	85	.13	.13	-.05	.07	.06
63	86	.08	.13	-.04	.13	.17
64	87	.04	.19	.19	.15	.32
65	88	.09	-.14	.01	.08	.06
66	90	.01	-.05	.33	.39	.49
67	91	.04	.02	.14	.37	.32
68	92	.08	.02	.10	.27	.20
69	93	.06	.14	.25	.35	.48
70	101	.09	.12	.25	.46	.54
71	102	.15	.23	.20	.19	.21
72	103	-.03	-.07	.18	.22	.34
73	104	.01	.19	.30	.17	.31
74	105	.15	.01	.09	.27	.14
75	107	.03	.10	.13	.24	.17
76	108	.19	.22	.27	.33	.28
77	109	.09	.10	.12	.24	.23
78	113	.14	.29	.38	.22	.43
79	115	.09	.10	.20	.36	.34
80	117	.07	.21	.10	.19	.22
81	119	-.01	.14	.11	-.01	.19
82	120	.12	.07	-.03	.24	.07
83	121	.07	.08	.09	.05	.13
84	122	.15	.16	.26	.40	.50
85	123	.17	.23	.30	.40	.39
86	126	.03	.27	.25	.21	.37
87	127	-.00	.16	.11	.28	.25
88	129	.11	.10	.25	.38	.37
89	131	.05	.09	.11	.16	.22
90	132	.01	.11	.01	.23	.12
91	135	.26	.22	.19	.15	.18
92	136	.17	.33	.34	.16	.29
93	137	.04	.21	.06	.03	.06
94	138	.16	.12	.10	.22	.21
95	140	.12	.09	.15	.17	.16
96	143	.30	.34	.22	.39	.28
97	144	.05	-.10	.21	.28	.27
98	145	.03	-.02	.03	.17	.17
99	146	.05	.12	.12	.15	.27
100	149	.16	.10	.20	.25	.21

VALIDITY STUDIES

The intent of the present study was to develop a series of paper-and-pencil measures of job performance against which aptitude tests such as the ASVAB could be validated. The validity study will, thus, primarily address the associations between the ASVAB scales and the three Criterion Tests. Specifically, validity assessment is developed in four phases:

1. The relationship between the ASVAB and the Criterion Tests.

2. The relationship between technical school grades (the previous criterion) and the ASVAB.

3. The relationship between technical school grades and the Criterion Tests.

4. Cross-validation studies.

Validation studies of this type usually concern the regression of an outcome measure on a test score. Phase 1 thus embodies the main features of the present validation effort. Phases 2 and 3 are included to assess the relationship of the previous criterion with the newly developed one. The last phase--cross-validation--is included to get an estimate of the degree of shrinkage in predictability involved with using weights derived from one sample to predict job performance in another independent group of subjects.

The Relationship of the ASVAB and the Criterion Tests

The means and standard deviations of each of the skill level groups on the nine ASVAB scales are presented in Table 16. A multivariate analysis of variance (MANOVA) performed on the ASVAB scores indicates that the skill level groups have significantly different mean profiles (asymptotic chi-square = 458.0319 with 27 degrees of freedom, $p < .0001$). The corresponding univariate tests indicate that two ASVAB subscales--Coding Speed and Automotive Information--are primarily responsible for this significant difference. From the mean scores presented in Table 16, it is noted that the basic airmen score higher than the older mechanics on Coding Speed, while scoring lower on Automotive Information. Coding Speed is a clerical skill heavily dependent upon quickness and mental agility. Thus, age logically could be a factor in these scores. The fact that the older, more skilled mechanics score higher than basic airmen on Automotive Information is also not surprising, given the mechanical exposure of these subjects.

TABLE 16 ASVAB MEANS AND STANDARD DEVIATIONS BY SKILL LEVEL

(Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

Skill Level		(1) Coding Speed	(2) Word Knowledge	(3) Arith. Reasoning	(4) Tool Knowledge	(5) Space Perception	(6) Mech. Comp.	(7) Shop Info.	(8) Auto. Info.	(9) Elec. Info.
		ASVAB Scale								
1	M	50.64	20.16	14.94	14.57	14.23	14.58	14.69	13.91	15.03
	SD	15.76	3.74	5.36	4.63	5.38	3.86	3.98	5.65	4.46
3	M	46.52	20.40	14.16	15.37	13.41	14.37	15.37	15.85	15.59
	SD	10.98	3.28	4.87	5.56	5.23	4.13	4.37	5.53	4.43
5	M	42.95	18.92	13.36	14.91	12.13	13.15	14.52	16.29	15.24
	SD	16.65	5.45	5.52	6.64	6.18	5.45	5.97	6.40	5.01
7	M	42.90	20.77	15.07	17.93	11.63	15.05	17.84	19.95	18.64
	SD	13.40	4.28	5.27	6.79	5.38	4.83	5.03	4.92	4.17

The correlation matrix for the nine ASVAB scales, pooled over skill levels, is presented in Table 17. Most of the correlations between ASVAB scales are in the low-to-moderate range, e.g., 30's to 40's. One exception to this general result involves the Information Scales (Shop, Automotive, and Electronic) where the correlations are noticeably higher. However, given the experience of the subjects involved, prior expectation is that the Information Scales will be highly correlated. The correlation matrix for level 1 personnel considered separately shows approximately the same pattern, except that the Information Scales are not as highly correlated.

TABLE 17 CORRELATIONS OF ASVAB SCALES POOLED OVER SKILL LEVELS

(Total N = 813; 3 = 194; 5 = 336; 7 = 283)

	1	2	3	4	5	6	7	8	9
1	1.00								
2	0.39	1.00							
3	0.39	0.50	1.00						
4	0.10	0.23	0.17	1.00					
5	0.32	0.36	0.44	0.39	1.00				
6	0.32	0.50	0.46	0.43	0.53	1.00			
7	0.22	0.49	0.34	0.61	0.37	0.60	1.00		
8	0.13	0.39	0.29	0.57	0.26	0.53	0.70	1.00	
9	0.18	0.49	0.38	0.49	0.31	0.53	0.68	0.65	1.00

Table 18 shows the correlations between individual ASVAB scales and the three Criterion Test scores plus the total test score. Individual correlation results are also presented by skill level.

None of the ASVAB scales is highly correlated with any of the Criterion Tests within any skill level. The individual ASVAB scales that correlate most highly with Criterion Test performance are the Information Scales (Shop, Automotive, and Electronic). Arithmetic Reasoning also has a slightly higher than average correlation with the Criterion Tests.

In order to more completely assess the relationship between the ASVAB scales and the Criterion Tests, a stepwise regression procedure was used to select the best sets of ASVAB predictors for the test scores. Stepwise regression is used to achieve a compromise between the increased precision gained by including as many predictors as possible, and the practical necessity of reducing complexity and costs by including only as many variables as necessary. A stepwise regression procedure enters into regression only those variables with an independent contribution to prediction that is statistically significant. After the first step in the process, the contribution of each variable already in regression is again

TABLE 18 CORRELATIONS OF ASVAB SCALES WITH CRITERION TEST SCORES BY SKILL LEVEL

(Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

Skill Level	Test	ASVAB Scale 1	2	3	4	5	6	7	8	9
1	1	.171	.273	.209	.247	.174	.250	.291	.303	.341
	2	.081	.216	.240	.218	.196	.178	.255	.225	.334
	3	.158	.211	.179	.095	.146	.131	.089	.079	.133
	T	.203	.349	.311	.286	.256	.283	.325	.313	.411
3	1	.231	.371	.283	.214	.257	.390	.360	.428	.399
	2	.010	.231	.201	.149	.264	.275	.114	.281	.275
	3	.147	.170	.236	.112	.282	.190	.110	.200	.255
	T	.179	.357	.308	.216	.329	.391	.289	.418	.413
5	1	.254	.293	.336	.116	.245	.228	.219	.299	.320
	2	.063	.159	.234	.064	.111	.134	.169	.211	.255
	3	.174	.240	.247	.095	.142	.127	.226	.168	.263
	T	.207	.283	.335	.112	.210	.205	.245	.283	.340
7	1	.245	.113	.208	.159	.190	.170	.091	.118	.165
	2	.197	.175	.229	.147	.193	.200	.178	.216	.235
	3	.217	.181	.269	.068	.152	.092	.092	.136	.155
	T	.258	.187	.279	.145	.209	.180	.143	.186	.218
3-5-7	1	.213	.291	.312	.211	.190	.281	.291	.365	.384
	2	.024	.189	.235	.210	.052	.204	.275	.362	.379
	3	.077	.214	.250	.201	.038	.165	.280	.323	.363
	T	.119	.265	.305	.238	.107	.249	.323	.402	.431

evaluated to see if it is needed, given that the new variable is now in regression. The process usually results in a regression equation containing those variables with the largest individual contribution to prediction.

Stepwise regression results using job performance test scores as criteria and the ASVAB scales as predictors are presented in Table 19. The results in Table 19 are given separately for pooled 3, 5, and 7 level personnel and for basic airmen (level 1). Basic airmen are separated from experienced jet mechanics for analysis purposes because of the possibility that the experience factor (e.g., time in service, attendance at technical school, job incumbency, and so forth) would moderate the relationship between the Criterion Tests and the ASVAB.

Separate regression solutions are also presented for the total sample and a "computing subsample." The results involving the computing subsample are to be used in the later cross-validation exercise. In all analyses, the three most predictive ASVAB scales are entered into regression.

TABLE 19 SUMMARY OF STEPWISE REGRESSION ANALYSES INVOLVING CRITERION TEST SCORES AND ASVAB SCALES

Test 1--Pooled 3, 5, and 7-Level Personnel (Total Sample; N = 813)

Variables in Regression--9, 3, and 8
Percentage of Variance Explained (R^2) = 18.8%
Regression Coefficients:
B0 = 37.07
B9 = 0.712
B3 = 0.525
B8 = 0.487

ANALYSIS OF VARIANCE

Source	DF	SS	MS	F	P
Total	812	244763.00			
Regression	3	45998.09	15332.70	62.41	.0001
Residual	809	198764.91	245.69		

Test 1--Pooled 3, 5, and 7-Level Personnel (Computing Subsample; N = 402)

Variables in Regression--9, 3, and 1
Percentage of Variance Explained = 18.0%
Regression Coefficients:
B0 = 38.40
B9 = 1.045
B3 = 0.385
B1 = 0.096

ANALYSIS OF VARIANCE

Source	DF	SS	MS	F	P
Total	401	108406.50			
Regression	3	19552.29	6517.43	29.19	.0001
Residual	398	88854.21	223.25		

Test 1--Basic Airmen (Level-1; N = 419)

Variables in Regression--9, 2, and 8
Percentage of Variance Explained: 17.1%
Regression Coefficients:
B0 = -11.878
B9 = 0.338
B2 = 0.383
B8 = 0.219

ANALYSIS OF VARIANCE

Source	DF	SS	MS	F	P
Total	418	23369.85			
Regression	3	4003.74	1334.58	28.60	.0001
Residual	415	19366.11	46.67		

TABLE 19 (continued)

Test 2--Pooled 3, 5, and 7-Level Personnel (Total Sample; N = 813)

Variables in Regression--9, 8, and 1
Percentage of Variance Explained: 17.4%
Regression Coefficients:
B0 = 8.120
B9 = 0.967
B8 = 0.632
B1 = -0.129

ANALYSIS OF VARIANCE

Source	DF	SS	MS	F	P
Total	812	259029.75			
Regression	3	45023.32	15007.77	56.73	.0001
Residual	809	214006.42	264.53		

Test 2--Pooled 3, 5, and 7-Level Personnel (Computing Subsample; N = 402)

Variables in Regression--9, 5, and 8
Percentage of Variance Explained: 18.4%
Regression Coefficients:
B0 = 8.884
B9 = 1.144
B5 = -0.538
B8 = 0.509

ANALYSIS OF VARIANCE

Source	DF	SS	MS	F	P
Total	401	124156.47			
Regression	3	22793.47	7597.82	29.83	.0001
Residual	398	101363.00	254.68		

Test 2--Basic Airmen (Level-1; N = 419)

Variables in Regression--9, 3, and 7
Percentage of Variance Explained: 14.1%
Regression Coefficients:
B0 = -5.073
B9 = 0.342
B3 = 0.187
B7 = 0.169

ANALYSIS OF VARIANCE

Source	DF	SS	MS	F	P
Total	418	17987.88			
Regression	3	2527.86	842.62	22.62	.0001
Residual	415	15460.02	37.25		

TABLE 19 (continued)

Test 3--Pooled 3, 5, and 7-Level Personnel (Total Sample; N = 813)

Variables in Regression--9, 8, and 6
Percentage of Variance Explained: 15.3%
Regression Coefficients:
B0 = 0.317
B9 = 1.096
B8 = 0.560
B6 = -0.469

ANALYSIS OF VARIANCE

Source	DF	SS	MS	F	P
Total	812	238692.72			
Regression	3	36527.64	12175.88	48.72	.0001
Residual	809	202165.08	249.90		

Test 3--Pooled 3, 5, and 7-Level Personnel (Computing Subsample; N = 402)

Variables in Regression--9, 3, and 5
Percentage of Variance Explained: 18.5%
Regression Coefficients:
B0 = -0.096
B9 = 1.240
B3 = 0.666
B5 = -0.608

ANALYSIS OF VARIANCE

Source	DF	SS	MS	F	P
Total	401	122331.89			
Regression	3	22633.00	7544.33	30.12	.0001
Residual	398	99698.88	250.50		

Test 3--Basic Airmen (Level-1; N = 419)

Variables in Regression--2, 1, and 4
Percentage of Variance Explained: 6.8%
Regression Coefficients:
B0 = -6.724
B2 = 0.294
B1 = 0.046
B4 = 0.108

ANALYSIS OF VARIANCE

Source	DF	SS	MS	F	P
Total	418	13965.04			
Regression	3	954.08	318.03	10.14	.0001
Residual	415	13010.96	31.35		

TABLE 19 (continued)

Total Test Score--Pooled 3, 5, and 7-Level Personnel (Total Sample; N = 813)

Variables in Regression--9, 8, and 3
Percentage of Variance Explained: 22.0%
Regression Coefficients:

B0 = 37.211
B9 = 2.371
B8 = 1.390
B3 = 1.073

ANALYSIS OF VARIANCE

Source	DF	SS	MS	F	P
Total	812	1689969.00			
Regression	3	371977.88	123992.63	76.11	.0001
Residual	809	1317991.12	1629.16		

Total Test Score--Pooled 3, 5, and 7-Level Personnel (Computing Subsample; N = 402)

Variables in Regression--9, 3, and 5
Percentage of Variance Explained: 22.6%
Regression Coefficients:

B0 = 49.011
B9 = 3.713
B3 = 1.610
B5 = -1.262

ANALYSIS OF VARIANCE

Source	DF	SS	MS	F	P
Total	401	815801.50			
Regression	3	184368.03	61456.01	38.74	.0001
Residual	398	631433.44	1586.52		

Total Test Score--Basic Airmen (Level-1; N = 419)

Variables in Regression--9, 2, and 4
Percentage of Variance Explained: 24.7%
Regression Coefficients:

B0 = -25.709
B9 = 0.786
B2 = 0.987
B4 = 0.442

ANALYSIS OF VARIANCE

Source	DF	SS	MS	F	P
Total	418	75095.65			
Regression	3	18573.60	6191.20	45.46	.0001
Residual	415	56522.05	136.20		

Table 19 presents a great deal of regression information. However, there are several patterns which appear to cut across the various analyses. The most obvious result is that all of the regression runs are highly significant. Next, considering the analyses involving the pooled 3, 5, and 7 level personnel, the most obvious result is that ASVAB scale 9--Electronic Information--is entered first in all the runs. The Electronics Information scale contains questions concerning knowledge of electricity and electronics. This scale thus appears to represent a high order experiential factor. ASVAB scale 8--Automotive Information--is also included in the final set of predictors in all analyses. Automotive Information does not, however, have the consistently high relationship with Criterion Test performance of Scale 9. The overall predictive efficiency of the various regression analyses as indexed by the squared multiple correlation (e.g., the percentage of variance explained) ranges from 15.3% for CT 3, up to 22.0% for the total test score. These squared multiple correlations are comparable to those usually obtained in studies involving the regression of aptitude or intelligence measures on achievement test performance.

The corresponding regression analyses involving the basic airmen indicate a similar pattern. In all cases except CT 3, Electronics Information is again the best single predictor of Criterion Test performance. The most potent predictor after Electronics Information is Word Knowledge. On CT 3, which primarily addresses the procedural and administrative aspects of job performance, Coding Speed and Word Knowledge are the best predictors of Criterion Test performance. However, the relationship between the ASVAB and CT 3 for basic airmen is substantially lower than is the case with the other tests.

In summary, the regression results involving experienced jet engine mechanics indicate that the ASVAB information scales, particulary Electronics Information and Automotive Information, are the most consistent predictors of Criterion Test performance. For basic airmen, Electronics Information is also a significant predictor; however, other ASVAB scales involving academic experience (e.g., Word Knowledge and Arithmetic Reasoning) are more important than is the case with the experienced mechanics. Taken together, the regression results suggest that previous mechanical or electronic experience plus general educational level, as indexed by verbal and arithmetic skills, are the best predictors of Criterion Test performance.

The above results are not surprising given the fact that many of the experienced mechanics taking the tests have several years of direct shop and mechanical experience. On-the-job experiences for experienced mechanics should result in related scores on the ASVAB information scales and the Criterion Tests. Additionally, verbal and arithmetic skills are usually a factor in test performance.

The ASVAB Versus Technical School Grades

The second phase of the validation study involves the relationship between technical school grades and the ASVAB. Phase 2, along with Phase 3, is included in the validation study for the following reasons:

1. To permit a comparison of relationships in the present sample with previously established relationships.

2. To obtain some information concerning the extent to which changing the criterion for selection will affect the composition of the pool of jet engine mechanics.

As previously indicated, the Air Force has traditionally used technical school grades (TSGs) as a criterion measure for the validation of aptitude tests. Changing the criterion measure could have the effect of selecting on different aptitude dimensions, thus changing the composition of the pool of basic airmen assigned to jet engine mechanic courses. In order to determine whether or not using the job knowledge tests as criteria will change the dimensions of selection, the ASVAB scales are regressed on TSGs, and the results compared with the previous finding concerning the relationships between the Criterion Tests and the ASVAB.

Table 20 presents the means and standard deviations of technical school grades separated by skill level. Since basic airmen have not been to the technical school, only experienced jet engine mechanics are included.

TABLE 20 MEANS AND STANDARD DEVIATIONS OF TECHNICAL SCHOOL GRADES BY SKILL LEVEL

	Skill Level		
	3	5	7
Mean	88.24	86.88	86.83
Standard Deviation	5.98	6.18	5.18
N	167 (86%)	275 (82%)	89 (31%)

The numbers of observations in each of the skill groups displayed in Table 20 are lower than the numbers indicated in previous analyses. This

decrement resulted from the fact that a number of subjects did not match the Technical Training Course Grade File, or matched the file but did not have a course number of 3ABR426X2 or 3ABR432X0. Several subjects matched the Technical Training Course Grade File and attended the course, but did not have a valid grade. The extent of subject loss, for one reason or another, is greatest for the 7 level personnel. Only 31% of the 7 level subjects have valid TSGs. The corresponding capture rates for level 3 and level 5 personnel are 86% and 82%, respectively.

The mean TSGs do not show much difference across skill levels. Mean TSG scores range from 88.24 to 86.83. An analysis of variance performed on the TSG mean scores across skill levels resulted in an F of 2.76, with 2 and 528 degrees of freedom; this value is not statistically significant ($p = .0625$). An additional observation concerning the TSG scores is the relatively small size of the standard deviations. The range of standard deviations is from 5.18 to 6.18. Such restricted ranges of TSG scores could have the effect of depressing the relationships of TSGs with the ASVAB scales and Criterion Tests.

The correlations of the ASVAB scales with TSGs, separated by skill level, are given in Table 21. None of the ASVAB subscales individually has a strong relationship with TSGs. The strongest associations, which hold across skill levels, are ASVAB scales 2, 3, 6, and 9--Word Knowledge, Arithmetic Reasoning, Mechanical Comprehension, and Electronics Information. The stepwise regression results summarized in Table 22 extend and clarify the correlation results.

TABLE 21 CORRELATIONS OF ASVAB SCALES WITH TECHNICAL SCHOOL GRADES BY SKILL LEVEL

(Total N = 531; 3 = 167; 5 = 275; 7 = 89)

		ASVAB Scale								
		1	2	3	4	5	6	7	8	9
Skill Level	3	.129	.407	.414	.121	.294	.347	.237	.228	.335
	5	.308	.339	.404	.092	.227	.277	.224	.283	.333
	7	.045	.082	.148	.171	.141	.437	.270	.286	.224
Pooled		.227	.321	.369	.109	.240	.317	.227	.241	.298

TABLE 22 SUMMARY OF STEPWISE REGRESSION RESULTS USING TECHNICAL SCHOOL GRADES AS THE CRITERION AND ASVAB SCALES AS PREDICTORS BY SKILL LEVEL

		Scales Included in Regression				Percentage of Variation Explained			
		Total Sample	N	Computing Subsample	N	Total Sample	N	Computing Subsample	N
Skill Level	3	3, 2, 9	167	3, 9, 1	96	28.1%	167	31.0%	96
	5	3, 9, 1	275	1, 3, 9	144	23.5%	275	27.2%	144
	7	6, 8, 2	89	6, 9, 3	42	20.5%	89	34.7%	42
Pooled		3, 9, 1	531	3, 1, 6	282	19.6%	531	25.1%	282

Considering Table 22, the variables entered into regression are similar for the 3 and 5 level personnel. ASVAB scales 3, 9, and either of scales 1 or 2, are selected in all cases. For level 7 personnel, the situation is different. Here, the regression results are less consistent, except for the fact that Mechanical Comprehension is always included in the set of predictors.

The results for the pooled sample represent a blending of the two previous conditions. In the pooled sample, Coding Speed and Word Knowledge are included in both regression analyses. However, Electronics Information is included in the runs involving the total sample, while Mechanical Comprehension is entered into regression in the runs involving the computing subsample.

The stepwise analyses summarized in Table 22 have somewhat higher levels of precision than is the case for the regression results involving the Criterion Tests and ASVAB scales. In the regressions involving the ASVAB and TSGs, the squared multiple correlations (SMCs) range from 19.6% to 34.7%. Additionally, the ASVAB scales most predictive of the dependent measure are different in the two situations. With the job-related Criterion Tests, ASVAB information measures are the best predictors. On the other hand, general academic skills such as Word Knowledge and Arithmetic Reasoning are most indicative of technical school performance. These latter scales are also related to Criterion Test performance, but not to the same extent as with TSGs. It thus appears that the two criteria--the job performance Criterion Tests and TSGs--are related to slightly different aspects of individual aptitude. Using the TSGs as a criterion would appear to favor better schooled, more academically proficient job candidates; using the Criterion Tests would apparently favor candidates having previous experience with, or interest in, mechanical/electrical tasks.

The Relationship Between the Criterion Tests and TSGs

In order to further clarify the relationship between the two criterion measures, the third phase of the validation study concerns the direct association between the Criterion Tests and technical school grades. Table 23 presents the correlations between the three Criterion Tests scores, their total, and TSGs.

TABLE 23 CORRELATIONS BETWEEN CRITERION TEST SCORES AND TECHNICAL SCHOOL GRADES BY SKILL LEVEL

		Criterion Test			
Skill Level	N	1	2	3	Total Score
3	167	.478	.423	.279	.570
5	275	.608	.381	.244	.525
7	89	.416	.277	.181	.346
Pooled	531	.528	.255	.134	.408

The correlations presented in Table 23 indicate several results worthy of mention. First, the correlations between TSGs and the Criterion Test scores are higher for CT 1 than the other tests. This result is not surprising, since CT 1 primarily addresses the job content of level 3 personnel. The material taught in technical school is most closely related to level 3 job skills. The job content items in CTs 2 and 3 do not relate as directly to the material presented in the technical school, thus their correlation with technical school performance should not be as high.

A second observation, concerning the relationship between Criterion Test performance and TSGs, is that, with one exception, the correlations decline with increasing skill level. This suggests that job incumbency reduces the relationship between school performance and the behaviors measured by the Criterion Tests. A variety of factors such as selection and on-the-job experience could account for this result.

The stepwise regression results summarized in Table 24 support and extend the previous correlation results. Table 24 presents only those variables which meet an entry requirement of having a partial-F statistic which is significant at the p = .05 level. If no variable is indicated, then no Criterion Test met the entry requirement.

TABLE 24 STEPWISE REGRESSION RESULTS USING TECHNICAL SCHOOL GRADES AS CRITERIA AND CRITERION TEST SCORES AS PREDICTORS

(Total N = 531; 3 = 167; 5 = 275; 7 = 89)

Skill Level	Variables in Regression				Percentage of Variation Explained			
Sample	Total	N	Computing	N	% Total	N	% Computing	N
3	1, 2	167	1, 3	96	28.2	167	37.9	96
5	1	275	1	144	36.9	275	31.6	144
7	1	89	-	42	17.3	89	-	42
Pooled	1, 3	531	1, 3	282	29.2	531	27.8	282

Not surprisingly, the regression results again indicate that scores on CT 1 are most predictive of technical school performance. CT 1 enters regression first in all cases. For skill levels 5 and 7, CT 1 is the only score significantly associated with technical school performance. The squared multiple correlations also indicate that the precision of prediction is higher for lower skill level mechanics than for the more experienced personnel.

Cross-Validation

The last phase of the validation study involves cross-validation of the regression results. In a restricted sense, cross-validation involves computing regression parameters in one sample of subjects and then assessing the predictive efficiency of these weights in a similar but independent group of subjects.

In order to accomplish the above objective, the total sample of subjects was, first, randomly divided into two subsamples denoted the "computing sample" and the "cross-validation sample." Regression parameters were derived using the computing sample, and then used to generate predicted scores in the cross-validation sample. In each case, weights were obtained for the three most predictive ASVAB scales. Before being used to generate predicted scores, the raw regression coefficients were converted to integer weights by multiplying by 10 and then rounding to the nearest integer. The final step involved computing the correlations between actual and predicted scores using the subjects in the cross-validation sample.

Table 25 presents the correlations computed between the actual and predicted scores for selected situations. Since only certain Criterion Tests are relevant for specific skill groups, e.g., CT 1 for skill level 3 and so forth, not all sets of results are of interest. Only those

TABLE 25 CORRELATIONS BETWEEN ACTUAL AND PREDICTED SCORES IN THE CROSS-VALIDATION SUBSAMPLE BY SKILL LEVEL

Skill Level	Variable															
	TSG				T1			T2			T3				Total Score		
	Cross-Val Sample		Comp Sample		Cross-Val Sample		Comp Sample	Cross-Val Sample		Comp Sample	Cross-Val Sample		Comp Sample		Cross-Val Sample		Comp Sample
Correlation:	r	r^2	R^2	N	r	r^2	R^2	r	r^2	R^2	r	r^2	R^2	N	r	r^2	R^2
3	.205	.042	.310	71	.372	.138	.416		N/A			N/A		98	.386	.149	.369
5	.455	.207	.272	131		N/A		.193	.037	.156		N/A		173	.306	.093	.216
7	.212	.045	.347	47		N/A			N/A		.232	.054	.084	140	.293	.086	.070
Pooled	.364	.132	.251	249	.420	.176	.180	.404	.163	.184	.370	.137	.185	411	.442	.195	.226

correlations relevant to a particular group are presented. For purposes of comparison and evaluation, the squared multiple correlations from the corresponding regression analysis involving the computing subsample are presented with the squared cross-validation correlation coefficients.

As is typically the case in studies of this nature, the squared correlations computed from the cross-validation sample show the shrinkage associated with applying regression weights in a second sample. The actual magnitude of the cross-validation correlations usually reflects the squared multiple correlations obtained in the computing subsample. However, in most cases, the squared correlations are so low that they must be considered practically not significant.

CONCLUSIONS

Summary

The purpose of the present study was to develop a set of job-related Criterion Tests against which to validate aptitude measures. Toward that objective, three 150-item job performance Criterion Tests were developed and administered to 1232 personnel ranging from basic airmen to 7-skill level senior jet engine mechanics. By selecting those items which best differentiated skill levels, and were most homogeneous in terms of their correlation with the total test score, the Criterion Tests were reduced to a final form of 100 items. Analysis of variance and correlation results indicated that scores on the Criterion Tests were highly associated with actual skill levels. Reliability assessment using coefficient alpha also demonstrated that subject responses were sufficiently consistent.

In terms of the associations between the Criterion Test scores and other measures, the results of the study indicated that ASVAB information subscales were most predictive of Criterion Test performance for experienced mechanics. Information measures were also related to test performance for basic airmen, but general knowledge subscales, e.g., Word Knowledge, played a more prominent role for these subjects.

When technical school grades were regressed on the ASVAB, academically oriented scales, such as Word Knowledge and Arithmetic Reasoning, were found to be most consistently predictive of technical school performance. A possible implication of this result is that changing the standard for selection as a jet engine mechanic from TSGs to the Criterion Tests may tend to select in favor of potential candidates with previous mechanical/electrical experience.

As a final step in comparing the two criterion measures, TSGs were regressed on the Criterion Tests. The results indicated that only CT 1 was consistently predictive of technical school performance. This result is not surprising, since CT 1 covers material similar to that presented in the technical school course of study.

The final phase of the study involved cross-validation. Here, regression results from a randomly selected half of the subjects were used to generate predicted scores for the other half. The derived correlations between the actual and predicted scores demonstrated the shrinkage typically found in regression cross-validation studies. Given the low percentages of variance accounted for in the initial regression analysis, the corresponding correlations in the cross-validation phase indicated a very low relationship between actual and predicted scores.

Discussion

Considered as a whole, the analytical results indicate a statistically significant relationship between Criterion Test performance and the aptitude measures. The degree of the relationship is such, however, that the practical significance of the relationship is marginal. That is, ASVAB scores cannot be used to reliably predict future job knowledge, as measured by the Criterion Tests. The Criterion Test scores do, however, appear to be quite indicative of the actual career level of jet engine mechanics.

The rather weak relationship between the Criterion Tests and the ASVAB found in the present study is typical of results obtained in other similar investigations. When aptitude measures are used to predict job success, the relationships are usually attenuated by a number of "moderator" variables. These moderator variables, such as age, IQ, geographical background, and so forth, serve to modify the relationship between the predictor variable and the criterion measures. One solution to this problem involves computing separate regression equations for major subgroups in the population. This could be done for low IQ and high IQ personnel, for urban versus rural personnel, and so forth. Another possibility, which possibly would involve less "capitalization on chance," is to include the moderator variables as predictors in the regression equation. This would be done by collecting a broad range of biographic data indicative of the background, relevant experiences, interests, and personality of the subjects, in addition to aptitude information. The biographical variables are then used, separately or in a combined form, in an attempt to account for variation that normally becomes part of residual error.

The present Criterion Test development study ideally lends itself to such a procedure. Not only are a sufficiently large number of subjects involved, but also a considerable amount of relevant biographical information has already been obtained on the subjects. However, due to the contractual limits of the study, the biographic variables were not included in the regression analysis. The consideration of biographical moderator variables thus remains an area for future study.

The current study also makes the assumption that career progression is a valid measure of job proficiency. If Criterion Test performance is associated with career progression, it is then assumed that the Criterion Tests may also be considered a measure of job proficiency. However, in reality, the validity of the first assumption is open to criticism. It might well be argued that career progression is really a function of endurance. That is, personnel who stay in the Air Force will get promoted. In that case, actual job proficiency may only marginally be a factor in promotion. In order to strengthen the theoretical bases of the present study, in the event the Criterion Tests are used to select personnel, it would be necessary to more completely specify the relationship between Criterion Test performance and actual job proficiency.

APPENDIX 1 ITEM ANALYSIS OF TEST 1 FOR 3-LEVEL AND 1-LEVEL PERSONNEL

(Total N = 613; 1 = 419; 3 = 194)

Rank	Item	$\phi(0,3)$	Diff(3)	$S^2(3)$	$r_{it(3)}$	Rank	Diff(0)	$S^2(0)$	$r_{it(0)}$	$\bar{R}$*	Include
1	81	.64	.93	.06	.195	105	.25	.19	-.002	53.0	Yes
2	62	.62	.71	.21	.235	86	.10	.09	.068	44.0	Yes
3	95	.62	.91	.08	.470	2	.24	.18	-.063	2.5	Yes
4	8	.61	.87	.11	.399	19	.22	.17	.003	11.5	Yes
5	35	.60	.79	.16	.153	122	.17	.14	.004	63.5	Yes
6	9	.59	.85	.13	.205	101	.22	.17	.079	53.5	Yes
7	77	.59	.88	.10	.277	62	.25	.19	.068	34.5	Yes
8	141	.59	.76	.18	.297	56	.16	.13	.151	32.0	Yes
9	10	.59	.81	.15	.274	64	.19	.15	-.055	36.5	Yes
10	90	.58	.86	.12	.369	32	.24	.18	-.010	21.0	Yes
11	36	.57	.74	.19	.169	117	.16	.13	-.093	64.0	Yes
12	126	.57	.81	.15	.346	39	.21	.17	.061	25.5	Yes
13	34	.56	.91	.08	.383	26	.32	.22	.194	19.5	Yes
14	23	.55	.64	.23	.370	30	.11	.10	.048	22.0	Yes
15	71	.55	.88	.10	.244	81	.29	.20	.086	48.0	Yes
16	41	.55	.85	.13	.325	48	.26	.19	.093	32.0	Yes
17	142	.55	.66	.22	.357	35	.12	.11	.121	26.0	Yes
18	17	.55	.86	.12	.220	97	.27	.20	.205	57.5	Yes
19	89	.55	.84	.14	.443	7	.25	.19	.067	13.0	Yes
20	99	.55	.89	.10	.414	14	.30	.21	-.009	17.0	Yes

*$\bar{R}$ is the average rank of $\phi(0,3)$ and $r_{it(3)}$.

APPENDIX 1 (continued)

Rank	Item	$\phi(0,3)$	Diff(3)	$S^2(3)$	$r_{it(3)}$	Rank	Diff(0)	$S^2(0)$	$r_{it(0)}$	$\bar{R}$*	Include
21	129	.54	.60	.24	.396	20	.10	.09	-.026	20.5	Yes
22	87	.53	.83	.14	.291	57	.27	.20	.050	39.5	Yes
23	52	.52	.61	.24	.342	43	.11	.10	.003	33.0	Yes
24	116	.52	.70	.21	.331	47	.17	.14	.167	35.5	Yes
25	75	.52	.71	.21	.344	42	.18	.15	.000	33.5	Yes
26	2	.51	.79	.17	.412	2	.24	.18	.078	14.0	Yes
27	69	.51	.79	.17	.413	15	.25	.19	.149	21.0	Yes
28	42	.50	.81	.15	.256	75	.28	.20	.008	51.5	Yes
29	32	.50	.80	.16	.214	99	.27	.20	-.003	64.0	Yes
30	54	.50	.93	.06	.139	127	.41	.24	.200	78.5	Yes
31	64	.49	.81	.15	.429	9	.29	.21	.094	20.0	Yes
32	146	.49	.56	.25	.247	77	.10	.09	.033	54.5	Yes
33	15	.49	.71	.21	.290	58	.21	.16	.151	45.5	Yes
34	120	.48	.74	.19	.384	25	.24	.18	.188	29.5	Yes
35	37	.47	.66	.22	.382	27	.18	.15	.082	31.0	Yes
36	48	.47	.89	.10	.106	135	.39	.24	.165	85.5	Yes
37	131	.47	.63	.23	.387	23	.16	.14	.045	30.0	Yes
38	3	.47	.77	.18	.258	73	.27	.20	.098	55.5	Yes
39	55	.47	.72	.20	.179	111	.23	.18	.003	75.0	Yes
40	112	.47	.72	.20	.356	36	.23	.18	.113	38.0	Yes

*$\bar{R}$ is the average rank of $\phi(0,3)$ and $r_{it(3)}$.

APPENDIX 1 (continued)

Rank	Item	$\phi(0,3)$	Diff(3)	$S^2(3)$	$r_{it(3)}$	Rank	Diff(0)	$S^2(0)$	$r_{it(0)}$	$\bar{R}$*	Include
41	91	.46	.93	.07	.234	89	.44	.25	.198	65.0	Yes
42	111	.46	.75	.19	.462	6	.27	.20	.129	24.0	Yes
43	59	.46	.89	.10	.363	33	.40	.24	.194	38.0	Yes
44	76	.45	.66	.22	.171	116	.20	.16	-.072	80.0	Yes
45	28	.44	.88	.10	.229	90	.41	.24	.118	67.5	Yes
46	136	.44	.46	.25	.386	24	.08	.07	.145	35.0	Yes
47	123	.44	.70	.21	.344	41	.24	.18	.138	44.0	Yes
48	72	.44	.90	.09	.436	8	.43	.25	.171	28.0	Yes
49	132	.43	.75	.19	.489	1	.30	.21	.220	25.0	Yes
50	82	.42	.89	.10	.178	112	.44	.25	.228	81.0	Yes
51	98	.42	.64	.23	.258	74	.21	.17	.081	62.5	Yes
52	104	.42	.79	.16	.404	18	.35	.23	.220	35.0	Yes
53	70	.42	.84	.14	.393	21	.39	.24	.165	37.0	Yes
54	102	.41	.87	.11	.176	114	.43	.25	.174	84.0	Yes
55	106	.41	.68	.22	.290	59	.26	.19	.014	57.0	Yes
56	27	.40	.54	.25	.168	118	.15	.13	.087	87.0	Yes
57	108	.40	.54	.25	.236	84	.15	.13	.025	70.5	Yes
58	65	.40	.76	.18	.417	13	.33	.22	.197	35.5	Yes
59	43	.40	.81	.15	.342	44	.38	.23	.088	51.5	Yes
60	61	.40	.66	.22	.324	49	.25	.19	.157	54.5	Yes

*$\bar{R}$ is the average rank of $\phi(0,3)$ and $r_{it(3)}$.

APPENDIX 1 (continued)

Rank	Item	$\phi(0,3)$	Diff(3)	$S^2(3)$	$r_{it}(3)$	Rank	Diff(0)	$S^2(0)$	$r_{it}(0)$	$\bar{R}$*	Include
61	18	.40	.95	.04	.152	123	.56	.25	.264	92.0	Yes
62	92	.39	.61	.24	.351	37	.21	.17	.020	49.5	Yes
63	117	.39	.77	.18	.208	100	.35	.23	-.061	81.5	Yes
64	21	.39	.91	.08	.271	67	.51	.25	.174	65.5	Yes
65	97	.38	.65	.23	.370	31	.26	.19	.024	48.0	Yes
66	134	.38	.59	.24	.218	98	.21	.17	.088	82.0	Yes
67	68	.37	.76	.18	.407	17	.36	.23	.181	42.0	Yes
68	58	.37	.74	.19	.240	83	.34	.23	.186	75.5	Yes
69	74	.37	.72	.20	.464	4	.32	.22	.040	36.5	Yes
70	63	.36	.96	.03	.175	115	.61	.24	.198	92.5	Yes
71	124	.36	.58	.24	.463	5	.22	.17	.140	38.0	Yes
72	83	.35	.81	.15	.235	85	.43	.25	.221	76.5	Yes
73	44	.35	.70	.21	.324	50	.32	.22	.143	61.5	Yes
74	46	.35	.75	.19	.245	79	.38	.23	.078	76.5	Yes
75	101	.35	.52	.25	.280	61	.18	.15	.044	68.0	Yes
76	94	.35	.65	.23	.338	46	.29	.21	.040	66.0	Yes
77	145	.34	.55	.25	.317	52	.20	.16	.221	64.5	Yes
78	40	.33	.95	.04	.203	103	.64	.23	.228	90.5	Yes
79	6	.33	.45	.25	.223	93	.15	.13	.001	86.0	Yes
80	49	.32	.71	.21	.304	54	.36	.23	.192	76.0	Yes

*$\bar{R}$ is the average rank of $\phi(0,3)$ and $r_{it}(3)$.

APPENDIX 1 (continued)

Rank	Item	$\phi(0,3)$	Diff(3)	$S^2(3)$	$r_{it(3)}$	Rank	Diff(0)	$S^2(0)$	$r_{it(0)}$	$\bar{R}$*	Include
81	150	.32	.66	.22	.361	34	.32	.22	.284	57.5	Yes
82	14	.32	.49	.25	.247	78	.18	.15	.088	80.0	Yes
83	121	.32	.53	.25	.393	22	.21	.17	.168	52.5	Yes
84	115	.31	.94	.06	.235	88	.65	.23	.213	86.0	Yes
85	93	.31	.59	.24	.224	92	.27	.20	.005	88.5	Yes
86	128	.30	.80	.16	.421	12	.49	.25	.028	49.0	Yes
87	22	.29	.35	.23	.118	131	.11	.10	-.046	109.0	No
88	110	.29	.47	.25	.338	45	.19	.15	-.013	66.5	Yes
89	96	.29	.86	.12	.467	3	.56	.25	.263	46.0	Yes
90	140	.28	.38	.24	.244	80	.13	.12	.121	85.0	Yes
91	73	.28	.45	.25	.195	106	.19	.15	.053	98.5	No
92	138	.28	.56	.25	.323	51	.27	.20	.287	71.5	Yes
93	125	.27	.32	.22	.190	107	.10	.09	.049	100.0	No
94	107	.27	.45	.25	.118	132	.19	.16	.046	113.0	No
95	31	.26	.36	.23	.269	69	.13	.11	.042	82.0	Yes
96	119	.26	.54	.25	.422	10	.27	.20	.218	53.0	Yes
97	149	.26	.48	.25	.262	71	.23	.18	.121	84.0	Yes
98	47	.26	.29	.21	.285	60	.09	.08	-.089	78.0	Yes
99	66	.25	.70	.21	.222	96	.42	.24	.191	97.5	Yes

*$\bar{R}$ is the average rank of $\phi(0,3)$ and $r_{it(3)}$.

APPENDIX 1 (continued)

Rank	Item	$\phi(0,3)$	Diff(3)	$S^2(3)$	$r_{it(3)}$	Rank	Diff(0)	$S^2(0)$	$r_{it(0)}$	$\overline{R}$*	Include
100	113	.25	.90	.09	.147	125	.66	.22	.189	112.5	No
101	13	.23	.57	.25	.188	109	.32	.22	.008	105.0	No
102	137	.23	.41	.24	.345	40	.19	.15	.128	71.0	Yes
103	79	.23	.48	.25	.125	129	.25	.19	.019	111.0	No
104	85	.23	.66	.22	.235	87	.41	.24	.118	95.5	No
105	135	.23	.47	.25	.272	65	.24	.18	.338	85.0	Yes
106	38	.23	.67	.22	.222	95	.42	.24	.253	100.5	No
107	78	.23	.35	23	.179	110	.15	.13	-.040	108.5	No
108	7	.22	.58	.24	.038	143	.34	.22	.097	125.5	No
109	50	.22	.77	.18	.110	133	.54	.25	.259	121.0	No
110	105	.22	.41	.24	.372	29	.20	.16	-.001	69.5	Yes
111	1	.21	.26	.19	.091	136	.10	.09	.008	123.5	No
112	5	.20	.42	.24	.275	63	.22	.17	.160	87.5	No
113	127	.20	.72	.20	.177	113	.51	.25	-.016	113.0	No
114	56	.20	.53	.25	.075	139	.32	.22	.083	126.5	No
115	67	.20	.42	.24	.189	108	.22	.17	.085	111.5	No
116	109	.19	.48	.25	.109	134	.28	.20	.052	125.0	No
117	26	.19	.60	.24	.377	28	.40	.24	.143	72.5	Yes
118	148	.19	.38	.23	.261	72	.20	.16	.116	95.0	No
119	133	.18	.42	.24	.167	119	.24	.18	.175	119.0	No

*$\overline{R}$ is the average rank of $\phi(0,3)$ and $r_{it(3)}$.

APPENDIX 1 (continued)

Rank	Item	$\phi(0,3)$	Diff(3)	$S^2(3)$	$r_{it(3)}$	Rank	Diff(0)	$S^2(0)$	$r_{it(0)}$	$\bar{R}$*	Include
120	100	.18	.46	.25	.270	68	.28	.20	.004	92.0	No
121	130	.17	.28	.20	.144	126	.14	.12	.148	123.5	No
122	86	.17	.35	.23	.271	66	.19	.15	.062	94.0	No
123	114	.17	.60	.24	.422	11	.42	.24	.218	67.0	Yes
124	122	.16	.76	.18	.201	104	.59	.24	.174	114.0	No
125	24	.14	.72	.20	.226	91	.58	.24	.215	108.0	No
126	16	.14	.45	.25	-.076	146	.31	.21	.057	136.0	No
127	20	.12	.53	.25	.135	128	.39	.24	-.003	127.5	No
128	4	.12	.79	.16	.348	38	.68	.22	.229	83.0	No
129	57	.12	.46	.25	.298	55	.34	.22	.025	92.0	No
130	84	.12	.39	.24	.204	102	.27	.20	.014	116.0	No
131	33	.10	.51	.25	.076	138	.39	.24	.225	134.5	No
132	139	.10	.17	.14	.048	141	.10	.09	-.052	136.5	No
133	25	.09	.55	.25	.254	76	.45	.25	.055	104.5	No
134	51	.05	.41	.24	.164	121	.35	.23	.275	127.5	No
135	39	.05	.43	.24	.066	140	.38	.23	.072	137.5	No
136	103	.04	.20	.16	-.037	144	.17	.14	-.022	140.0	No
137	144	.03	.18	.14	.085	137	.15	.13	.115	137.0	No
138	143	.03	.60	.24	.166	120	.57	.25	.040	129.0	No
139	60	.02	.38	.23	.243	82	.36	.23	.129	110.5	No

*$\bar{R}$ is the average rank of $\phi(0,3)$ and $r_{it(3)}$.

APPENDIX 1 (continued)

Rank	Item	$\phi(0,3)$	Diff(3)	$S^2(3)$	$r_{it(3)}$	Rank	Diff(0)	$S^2(0)$	$r_{it(0)}$	$\bar{R}$*	Include
140	11	.01	.65	.23	.314	53	.63	.23	.240	96.5	No
141	147	.01	.59	.24	.120	130	.58	.24	-.044	135.5	No
142	29	-.01	.19	.15	-.098	148	.20	.16	-.006	145.0	No
143	118	-.01	.53	.25	.223	94	.54	.25	.223	118.5	No
144	45	-.02	.49	.25	.264	70	.51	.25	.079	107.0	No
145	19	-.04	.36	.23	.039	142	.39	.24	.073	143.5	No
146	30	-.08	.22	.17	-.044	145	.30	.21	.040	145.5	No
147	88	-.08	.22	.17	.148	124	.30	.21	.152	135.5	No
148	53	-.11	.09	.08	-.078	147	.17	.14	.060	147.5	No
149	80	-.19	.09	.08	-.139	149	.26	.19	.033	149.0	No
150	12	-.36	.19	.15	-.176	150	.58	.24	.121	150.0	No

*$\bar{R}$ is the average rank of $\phi(0,3)$ and $r_{it(3)}$.

APPENDIX 2 ITEM ANALYSIS OF TEST 2 FOR 5-LEVEL and 3-LEVEL PERSONNEL

(Total N = 530; 3 = 194; 5 = 336)

Rank	Item	$\phi(3,5)$	Diff(5)	$S^2(5)$	$r_{it(5)}$	Rank	Diff(3)	$S^2(3)$	$r_{it(3)}$	$\overline{R}$*	Include
1	80	.21	.77	.18	.403	4	.57	.24	.235	2.5	Yes
2	138	.20	.49	.25	.334	15	.28	.20	.191	8.5	Yes
3	65	.19	.43	.25	.197	63	.24	.18	.220	33.0	Yes
4	131	.16	.33	.22	.264	41	.19	.15	-.027	22.5	Yes
5	123	.15	.42	.24	.290	29	.27	.20	.195	17.0	Yes
6	104	.15	.40	.24	.368	10	.26	.19	.180	8.0	Yes
7	99	.14	.55	.25	.301	23	.40	.24	.208	15.0	Yes
8	9	.14	.69	.21	.318	19	.55	.25	.283	13.5	Yes
9	3	.14	.31	.21	.211	56	.19	.15	.039	32.5	Yes
10	43	.14	.75	.19	.384	5	.62	.24	.399	7.5	Yes
11	145	.13	.35	.23	.187	70	.23	.18	.149	40.5	Yes
12	129	.13	.50	.25	.188	69	.36	.23	.187	40.5	Yes
13	62	.13	.63	.23	.298	24	.49	.25	.365	18.5	Yes
14	60	.13	.61	.24	.351	12	.48	.25	.254	13.0	Yes
15	22	.13	.68	.22	.330	16	.55	.25	.310	15.5	Yes
16	27	.12	.19	.15	.132	88	.10	.09	.049	47.0	Yes
17	124	.12	.67	.22	.159	80	.55	.25	.151	48.5	Yes
18	147	.12	.35	.23	.112	97	.23	.18	.077	57.5	Yes
19	78	.11	.43	.25	.328	17	.32	.22	.159	18.0	Yes
20	50	.11	.85	.13	.424	2	.76	.18	.371	11.0	Yes

*$\overline{R}$ is the average rank of $\phi(3,5)$ and $r_{it(5)}$.

APPENDIX 2 (continued)

Rank	Item	$\phi(3,5)$	Diff(5)	$S^2(5)$	$r_{it}(5)$	Rank	Diff(3)	$S^2(3)$	$r_{it}(3)$	$\overline{R}$*	Include
21	39	.11	.63	.23	.223	51	.52	.25	.279	36.0	Yes
22	64	.11	.40	.24	.190	66	.29	.21	.140	44.0	Yes
23	116	.11	.62	.24	.316	20	.51	.25	.263	21.5	Yes
24	24	.10	.36	.23	.181	74	.26	.19	.311	59.0	Yes
25	57	.10	.34	.22	.128	92	.24	.18	.032	58.5	Yes
26	130	.10	.33	.22	.421	3	.23	.18	.279	14.5	Yes
27	16	.09	.19	.15	.159	81	.12	.10	-.021	54.0	Yes
28	92	.09	.36	.23	.033	125	.27	.20	.037	76.5	Yes
29	14	.09	.28	.20	.212	55	.20	.16	.268	42.0	Yes
30	25	.09	.65	.23	.384	6	.56	.25	.338	18.0	Yes
31	51	.09	.55	.25	.209	57	.46	.25	.325	44.0	Yes
32	45	.08	.12	.10	.063	114	.07	.06	-.124	73.0	Yes
33	34	.08	.39	.24	.122	93	.30	.21	-.025	63.0	Yes
34	18	.08	.38	.24	.197	64	.30	.21	.225	49.0	Yes
35	48	.08	.43	.25	.172	77	.35	.23	.119	56.0	Yes
36	33	.08	.74	.19	.383	8	.66	.22	.314	22.0	Yes
37	12	.08	.53	.25	.178	76	.45	.25	.206	56.5	Yes
38	121	.07	.38	.24	.189	67	.31	.21	.172	52.5	Yes
39	142	.07	.49	.25	.186	71	.41	.24	.206	55.0	Yes
40	5	.07	.46	.25	.287	31	.39	.24	.114	35.5	Yes

* $\overline{R}$ is the average rank of $\phi(3,5)$ and $r_{it}(5)$.

APPENDIX 2 (continued)

Rank	Item	$\phi(3,5)$	Diff(5)	$S^2(5)$	$r_{it(5)}$	Rank	Diff(3)	$S^2(3)$	$r_{it(3)}$	$\bar{R}$*	Include
41	112	.07	.53	.25	.266	39	.46	.25	.332	40.0	Yes
42	37	.06	.82	.15	.376	9	.77	.18	.366	25.5	Yes
43	84	.06	.21	.16	.157	82	.15	.13	.090	62.5	Yes
44	71	.06	.28	.20	.167	78	.23	.18	-.131	61.0	Yes
45	82	.06	.73	.20	.279	35	.68	.22	.251	40.0	Yes
46	109	.06	.44	.25	.234	50	.38	.23	.285	48.0	Yes
47	2	.06	.35	.23	.196	65	.29	.21	.261	56.0	Yes
48	13	.06	.45	.25	.134	87	.39	.24	.186	67.5	Yes
49	83	.06	.66	.22	.252	43	.60	.24	.318	46.0	Yes
50	68	.06	.22	.17	.074	110	.17	.14	-.041	80.0	Yes
51	107	.06	.32	.22	-.030	139	.27	.20	.166	95.0	Yes
52	134	.06	.66	.22	.296	26	.61	.24	.334	39.0	Yes
53	103	.05	.42	.24	.238	48	.36	.23	.061	50.5	Yes
54	133	.05	.20	.16	-.012	135	.16	.13	.013	94.5	Yes
55	106	.05	.35	.23	.113	95	.30	.21	.228	75.0	Yes
56	46	.05	.24	.18	.161	79	.20	.16	.032	67.5	Yes
57	42	.05	.33	.22	.221	52	.28	.20	.135	54.5	Yes
58	7	.04	.10	.09	-.026	138	.08	.07	-.006	98.0	Yes
59	38	.04	.51	.25	.287	32	.46	.25	.377	45.5	Yes

*$\bar{R}$ is the average rank of $\phi(3,5)$ and $r_{it(5)}$.

APPENDIX 2 (continued)

Rank	Item	$\phi(3,5)$	Diff(5)	$S^2(5)$	$r_{it(5)}$	Rank	Diff(3)	$S^2(3)$	$r_{it(3)}$	$\overline{R}$*	Include
60	55	.04	.43	.24	.198	62	.38	.24	.129	61.0	Yes
61	144	.04	.21	.16	.091	103	.17	.14	-.050	82.0	Yes
62	105	.04	.42	.24	.085	105	.38	.23	.137	83.5	Yes
63	135	.04	.55	.25	.311	22	.51	.25	.273	42.5	Yes
64	88	.04	.59	.24	.274	38	.55	.25	.347	51.0	Yes
65	74	.04	.43	.24	.181	73	.39	.24	.240	69.0	Yes
66	127	.04	.35	.23	.054	118	.31	.22	.106	92.0	Yes
67	47	.04	.21	.16	.136	84	.18	.14	.002	75.5	Yes
68	59	.04	.63	.23	.287	30	.59	.24	.403	49.0	Yes
69	93	.04	.35	.23	.238	47	.31	.21	.278	58.0	Yes
70	31	.04	.67	.22	.216	53	.63	.23	.282	61.5	Yes
71	66	.04	.19	.15	-.007	133	.16	.13	.005	102.0	Yes
72	149	.04	.40	.24	.252	44	.37	.23	.225	58.0	Yes
73	19	.04	.60	.24	.336	14	.56	.25	.252	43.5	Yes
74	69	.04	.46	.25	.343	13	.42	.24	.329	43.5	Yes
75	52	.04	.21	.17	.078	108	.18	.15	.082	91.5	Yes
76	118	.04	.44	.25	.205	58	.41	.24	.110	67.0	Yes
77	4	.03	.32	.22	.060	116	.28	.20	-.053	96.5	Yes
78	75	.03	.24	.18	.101	100	.21	.16	-.103	89.0	Yes
79	35	.03	.12	.11	-.059	141	.10	.09	-.033	60.0	Yes

*$\overline{R}$ is the average rank of $\phi(3,5)$ and $r_{it(5)}$.

APPENDIX 2 (continued)

Rank	Item	$\phi(3,5)$	Diff(5)	$S^2(5)$	$r_{it(5)}$	Rank	Diff(3)	$S^2(3)$	$r_{it(3)}$	$\bar{R}$*	Include
80	85	.03	.20	.16	.041	122	.17	.14	.109	101.0	Yes
81	136	.03	.20	.16	.202	60	.17	.14	.071	70.5	Yes
82	90	.03	.57	.24	.384	7	.54	.25	.373	44.5	Yes
83	20	.03	.21	.17	.111	98	.19	.15	.137	90.5	Yes
84	26	.03	.29	.21	.057	117	.26	.19	.218	100.5	Yes
85	108	.03	.35	.23	.082	106	.32	.22	.113	95.5	Yes
86	139	.03	.30	.21	.043	121	.27	.20	.081	103.5	Yes
87	32	.03	.07	.06	-.015	136	.05	.05	-.181	111.5	Yes
88	89	.03	.63	.23	.297	25	.61	.24	.432	106.5	Yes
89	86	.02	.61	.24	.265	40	.58	.24	.404	64.5	Yes
90	111	.02	.28	.20	.180	75	.26	.19	.264	82.5	Yes
91	101	.02	.18	.14	.113	96	.16	.13	-.143	93.5	Yes
92	76	.02	.83	.14	.434	1	.81	.15	.419	46.5	Yes
93	117	.02	.34	.22	.105	99	.32	.22	.094	96.5	Yes
94	77	.02	.36	.23	.054	119	.34	.22	.074	106.5	Yes
95	97	.02	.16	.14	-.281	150	.15	.13	-.160	122.5	No
96	119	.02	.20	.16	.010	130	.19	.15	.002	113.0	Yes
97	67	.01	.71	.21	.246	45	.70	.21	.301	71.0	Yes
98	125	.01	.33	.22	-.043	140	.32	.22	-.011	119.0	No
99	114	.01	.64	.23	.284	33	.63	.23	.252	66.0	Yes

* $\bar{R}$ is the average rank of $\phi(3,5)$ and $r_{it(5)}$.

APPENDIX 2 (continued)

Rank	Item	$\phi(3,5)$	Diff(5)	$S^2(5)$	$r_{it(5)}$	Rank	Diff(3)	$S^2(3)$	$r_{it(3)}$	$\bar{R}$*	Include
100	30	.01	.63	.23	.234	49	.62	.24	.288	74.5	Yes
101	29	.01	.59	.24	.202	59	.58	.24	.169	80.0	Yes
102	100	.01	.34	.23	.184	72	.34	.22	.144	87.0	Yes
103	54	.00	.30	.21	.096	101	.29	.21	.152	102.0	No
104	128	.00	.19	.16	.034	124	.19	.15	-.006	114.0	No
105	79	.00	.19	.15	-.085	144	.19	.15	-.015	124.5	No
106	49	.00	.82	.15	.283	34	.82	.15	.286	70.0	No
107	53	.00	.26	.19	-.079	143	.26	.19	-.109	125.0	No
108	28	.00	.21	.17	.039	123	.21	.17	.085	115.5	No
109	120	.00	.30	.21	.147	83	.30	.21	.086	96.0	No
110	94	.00	.24	.18	-.093	146	.25	.19	.078	128.0	No
111	63	.00	.29	.21	.094	102	.30	.21	.051	106.5	No
112	98	-.01	.17	.14	-.070	142	.18	.14	-.089	127.0	No
113	81	-.01	.71	.20	.368	11	.72	.20	.236	62.0	No
114	95	-.01	.53	.25	.321	18	.54	.25	.192	66.0	No
115	146	-.01	.40	.24	.131	89	.41	.24	.142	102.0	No
116	140	-.01	.41	.24	.214	54	.42	.24	.269	80.0	No
117	61	-.01	.47	.25	.022	127	.48	.25	.162	122.0	No
118	40	-.01	.85	.12	.294	27	.86	.12	.275	72.5	No
119	23	-.01	.21	.16	.024	126	.22	.17	.187	121.5	No

*$\bar{R}$ is the average rank of $\phi(3,5)$ and $r_{it(5)}$.

APPENDIX 2 (continued)

Rank	Item	$\phi(3,5)$	Diff(5)	$S^2(5)$	$r_{it(5)}$	Rank	Diff(3)	$S^2(3)$	$r_{it(3)}$	$\overline{R}$*	Include
120	137	-.02	.31	.21	.134	85	.32	.22	.040	102.5	No
121	1	-.02	.29	.21	-.087	145	.30	.21	.055	133.0	No
122	58	-.02	.63	.23	.291	28	.65	.23	.471	75.0	No
123	73	-.02	.45	.25	.255	42	.47	.25	.349	82.5	No
124	11	-.02	.22	.17	.075	109	.24	.18	.094	116.5	No
125	17	-.03	.39	.24	.063	113	.41	.24	.102	119.0	No
126	143	-.03	.36	.23	.130	91	.39	.24	.080	108.5	No
127	126	-.03	.24	.18	.072	112	.27	.20	-.081	118.5	No
128	41	-.03	.32	.22	.062	115	.35	.23	.004	121.5	No
129	70	-.03	.21	.16	.131	90	.23	.18	.058	109.5	No
130	110	-.03	.22	.17	.079	107	.25	.19	.167	118.5	No
131	6	-.03	.06	.06	-.007	132	.08	.07	.093	131.5	No
132	132	-.04	.14	.12	-.131	149	.16	.14	-.129	140.5	No
133	87	-.04	.35	.23	.275	37	.39	.24	.222	85.0	No
134	148	-.04	.16	.13	-.007	134	.20	.16	-.068	134.0	No
135	56	-.05	.47	.25	.315	21	.52	.25	.345	78.0	No
136	113	-.05	.35	.23	.241	46	.40	.24	.187	91.0	No
137	122	-.05	.49	.25	.201	61	.54	.25	.172	99.0	No
138	96	-.05	.37	.23	.188	68	.43	.24	.100	103.0	No
139	10	-.06	.12	.11	.008	131	.16	.14	.003	85.0	No

*$\overline{R}$ is the average rank of $\phi(3,5)$ and $r_{it(5)}$.

APPENDIX 2 (continued)

Rank	Item	$\phi(3,5)$	Diff(5)	$S^2(5)$	$r_{it(5)}$	Rank	Diff(3)	$S^2(3)$	$r_{it(3)}$	$\bar{R}$*	Include
140	102	-.06	.33	.22	.134	86	.39	.24	.256	113.0	No
141	44	-.06	.22	.17	.014	129	.28	.20	.035	135.0	No
142	72	-.06	.12	.11	-.129	148	.17	.14	-.115	145.0	No
143	15	-.07	.13	.11	.073	111	.18	.15	-.127	127.5	No
144	21	-.07	.24	.18	.016	128	.31	.21	.086	136.5	No
145	141	-.08	.12	.10	-.104	147	.17	.14	-.117	146.0	No
146	36	-.08	.25	.19	.086	104	.32	.22	.107	125.0	No
147	150	-.08	.22	.17	.120	94	.29	.21	.085	120.5	No
148	115	-.13	.12	.11	-.017	137	.22	.17	.051	142.5	No
149	8	-.18	.43	.24	.277	36	.61	.24	.272	92.5	No
150	91	-.23	.24	.18	.053	120	.46	.25	.153	135.0	No

* $\bar{R}$ is the average rank of $\phi(3,5)$ and $r_{it(5)}$.

APPENDIX 3 ITEM ANALYSIS OF TEST 3 FOR 7-LEVEL AND 5-LEVEL PERSONNEL

(Total N = 619; 5 = 336; 7 = 283)

Rank	Item	$\phi(5,7)$	Diff(7)	$S^2(7)$	$r_{it(7)}$	Rank	Diff(5)	$S^2(5)$	$r_{it(5)}$	$\bar{R}$*	Include
1	101	.48	.74	.19	.434	4	.26	.19	.235	2.5	Yes
2	25	.48	.66	.22	.321	23	.19	.15	.227	12.5	Yes
3	18	.47	.71	.21	.228	51	.24	.18	.355	27.0	Yes
4	35	.43	.65	.23	.340	18	.22	.17	.326	11.0	Yes
5	57	.43	.51	.25	.529	1	.12	.10	.243	3.0	Yes
6	122	.40	.79	.17	.377	7	.39	.24	.229	6.5	Yes
7	45	.39	.94	.06	.245	42	.60	.24	.277	24.5	Yes
8	93	.38	.69	.22	.364	8	.30	.21	.219	8.0	Yes
9	90	.38	.65	.23	.385	6	.27	.20	.260	7.5	Yes
10	17	.38	.71	.21	.258	36	.33	.22	.260	23.0	Yes
11	47	.36	.46	.25	.305	26	.13	.12	.082	18.5	Yes
12	2	.34	.87	.11	.280	31	.57	.25	.281	21.5	Yes
13	40	.33	.40	.24	.350	13	.11	.10	.208	13.0	Yes
14	19	.32	.48	.25	.160	76	.18	.15	.154	45.0	Yes
15	44	.32	.68	.22	.447	2	.36	.23	.321	8.5	Yes
16	39	.31	.73	.20	.286	29	.42	.24	.202	22.5	Yes
17	113	.30	.72	.20	.199	59	.43	.24	.391	38.0	Yes
18	81	.30	.89	.10	.255	38	.63	.23	.243	28.0	Yes
19	31	.29	.43	.25	.280	32	.17	.14	.127	25.5	Yes
20	103	.29	.55	.25	.195	61	.27	.20	.156	40.5	Yes

* $\bar{R}$ is the average rank of $\phi(5,7)$ and $r_{it(7)}$.

APPENDIX 3 (continued)

Rank	Item	$\phi(5,7)$	Diff(7)	$S^2(7)$	$r_{it(7)}$	Rank	Diff(5)	$S^2(5)$	$r_{it(5)}$	$\bar{R}$*	Include
21	80	.28	.58	.24	.412	5	.30	.21	.210	13.0	Yes
22	43	.28	.72	.20	.146	80	.44	.25	.199	51.0	Yes
23	72	.27	.46	.25	.251	39	.20	.16	.249	31.0	Yes
24	115	.27	.51	.25	.340	17	.25	.19	.205	20.5	Yes
25	62	.27	.78	.17	.311	24	.53	.25	.266	24.5	Yes
26	20	.26	.94	.06	.244	44	.74	.19	.251	35.0	Yes
27	91	.24	.68	.22	.352	12	.44	.25	.140	19.5	Yes
28	87	.24	.59	.24	.129	87	.35	.23	.166	57.5	Yes
29	126	.23	.78	.17	.202	57	.56	.25	.234	43.0	Yes
30	22	.23	.46	.25	.356	11	.24	.18	.171	20.5	Yes
31	29	.23	.90	.09	.291	28	.72	.20	.329	29.5	Yes
32	68	.23	.80	.16	.272	34	.59	.24	.284	33.0	Yes
33	53	.22	.37	.23	.329	21	.17	.14	.098	27.0	Yes
34	11	.22	.47	.25	.189	63	.26	.19	.225	48.5	Yes
35	146	.22	.62	.24	.134	84	.40	.24	.094	59.5	Yes
36	63	.22	.59	.24	.216	54	.38	.24	.182	45.0	Yes
37	8	.21	.60	.24	.272	33	.39	.24	.123	35.0	Yes
38	123	.21	.83	.14	.360	9	.64	.23	.283	23.5	Yes
39	73	.21	.71	.20	.342	16	.51	.25	.189	27.5	Yes
40	21	.21	.58	.24	.348	14	.37	.23	.206	27.0	Yes

*$\bar{R}$ is the average rank of $\phi(5,7)$ and $r_{it(7)}$.

APPENDIX 3 (continued)

Rank	Item	$\phi(5,7)$	Diff(7)	$S^2(7)$	$r_{it(7)}$	Rank	Diff(5)	$S^2(5)$	$r_{it(5)}$	$\bar{R}$*	Include
41	14	.21	.64	.23	.337	19	.43	.24	.320	30.0	Yes
42	119	.21	.54	.25	-.022	133	.34	.22	.102	87.5	Yes
43	129	.21	.65	.23	.344	15	.44	.25	.243	29.5	Yes
44	75	.20	.51	.25	.189	62	.32	.22	.205	53.0	Yes
45	144	.19	.27	.20	.243	46	.12	.10	.195	45.5	Yes
46	41	.18	.51	.25	.299	27	.33	.22	.230	36.5	Yes
47	131	.18	.49	.25	.185	67	.32	.22	.116	57.0	Yes
48	27	.18	.98	.02	.183	68	.88	.11	.264	58.0	Yes
49	10	.18	.19	.15	.041	116	.07	.07	-.062	82.5	Yes
50	13	.18	.24	.18	.156	78	.10	.09	.106	64.0	Yes
51	83	.18	.49	.25	.213	55	.32	.22	.212	53.0	Yes
52	145	.17	.30	.21	.144	81	.15	.13	.024	66.5	Yes
53	136	.17	.93	.07	.176	71	.82	.15	.323	62.0	Yes
54	109	.17	.40	.24	.243	45	.24	.18	.147	49.5	Yes
55	64	.16	.69	.21	.438	3	.53	.25	.196	29.0	Yes
56	104	.16	.55	.25	.171	73	.39	.24	.290	64.5	Yes
57	69	.15	.42	.24	.163	75	.28	.20	.192	66.0	Yes
58	86	.15	.45	.25	.106	95	.31	.21	-.069	76.5	Yes
59	127	.15	.41	.24	.281	30	.26	.19	.139	44.5	Yes
60	66	.15	.77	.18	.244	43	.64	.23	.388	51.5	Yes

*$\bar{R}$ is the average rank of $\phi(5,7)$ and $r_{it(7)}$.

APPENDIX 3 (continued)

Rank	Item	$\phi(5,7)$	Diff(7)	$S^2(7)$	$r_{it(7)}$	Rank	Diff(5)	$S^2(5)$	$r_{it(5)}$	$\bar{R}$*	Include
61	12	.14	.28	.20	.131	86	.16	.13	.078	73.5	Yes
62	51	.14	.81	.16	.221	52	.68	.22	.138	57.0	Yes
63	117	.14	.29	.21	.189	64	.17	.14	.082	63.5	Yes
64	28	.14	.62	.24	.186	66	.48	.25	.116	65.0	Yes
65	92	.13	.33	.22	.270	35	.21	.17	.126	50.0	Yes
66	56	.13	.70	.21	.245	41	.58	.24	.203	53.5	Yes
67	4	.13	.47	.25	.142	83	.35	.23	166	75.0	Yes
68	59	.12	.67	.22	.118	92	.55	.25	.173	80.0	Yes
69	49	.12	.35	.23	.177	69	.24	.18	.046	69.0	Yes
70	55	.12	.89	.10	.332	20	.81	.16	.321	45.0	Yes
71	121	.12	.45	.25	.042	115	.34	.22	.092	93.0	Yes
72	16	.12	.34	.22	.001	130	.24	.18	.066	101.0	Yes
73	6	.11	.61	.24	.307	25	.50	.25	.195	49.0	Yes
74	52	.11	.40	.24	.060	108	.29	.21	.116	91.0	Yes
75	77	.11	.23	.18	.052	109	.15	.12	.069	92.0	Yes
76	54	.11	.20	.16	.049	111	.12	.11	.138	93.5	Yes
77	149	.10	.65	.23	.204	56	.54	.25	.195	66.5	Yes
78	138	.10	.44	.25	.196	60	.34	.22	.103	69.0	Yes
79	102	.09	.58	.24	.200	58	.49	.25	.187	68.5	Yes
80	107	.09	.30	.21	.230	50	.22	.17	.118	65.0	Yes

*$\bar{R}$ is the average rank of $\phi(5,7)$ and $r_{it(7)}$.

APPENDIX 3 (continued)

Rank	Item	$\phi(5,7)$	Diff(7)	$S^2(7)$	$r_{it(7)}$	Rank	Diff(5)	$S^2(5)$	$r_{it(5)}$	$\overline{R}$*	Include
81	140	.09	.34	.23	.164	74	.26	.19	.124	77.5	Yes
82	143	.09	.75	.19	.358	10	.67	.22	.220	46.0	Yes
83	108	.08	.65	.23	.327	22	.57	.25	.262	52.5	Yes
84	79	.08	.48	.25	.219	53	.40	.24	.279	68.5	Yes
85	34	.08	.35	.23	.106	96	.27	.20	.115	90.5	Yes
86	82	.07	.44	.25	.176	70	.37	.23	.264	78.0	Yes
87	141	.07	.43	.25	-.024	134	.36	.23	.128	110.5	No
88	15	.07	.17	.14	.127	88	.12	.10	-.077	88.0	Yes
89	85	.07	.24	.18	.090	100	.18	.15	-.034	92.5	Yes
90	24	.07	.64	.23	.257	37	.57	.24	.084	63.5	Yes
91	137	.07	.29	.21	.042	114	.23	.18	.037	102.5	Yes
92	32	.07	.55	.25	.106	97	.48	.25	.150	94.5	Yes
93	9	.07	.27	.20	.188	65	.21	.17	.007	79.0	Yes
94	88	.07	.16	.13	.103	99	.11	.10	.023	96.5	Yes
95	46	.06	.12	.11	.133	85	.09	.08	.026	90.0	Yes
96	110	.06	.15	.13	-.024	135	.11	.10	-.020	115.5	No
97	114	.05	.64	.23	.011	126	.58	.24	.179	111.5	No
98	3	.05	.39	.24	.050	110	.34	.22	.151	104.0	No
99	132	.04	.40	.24	.246	40	.35	.23	.051	69.5	Yes
100	128	.04	.23	.18	-.086	145	.20	.16	.071	122.5	No

* $\overline{R}$ is the average rank of $\phi(5,7)$ and $r_{it(7)}$.

APPENDIX 3 (continued)

Rank	Item	$\phi(5,7)$	Diff(7)	$S^2(7)$	$r_{it(7)}$	Rank	Diff(5)	$S^2(5)$	$r_{it(5)}$	$\overline{R}$*	Include
101	23	.04	.23	.18	.020	122	.20	.16	.100	111.5	No
102	135	.03	.46	.25	.142	82	.43	.24	.181	92.0	Yes
103	70	.03	.18	.15	.078	104	.16	.13	.043	103.5	No
104	112	.03	.32	.22	.033	118	.29	.21	.094	111.0	No
105	84	.02	.18	.15	.015	124	.17	.14	-.125	114.5	No
106	105	.02	.51	.25	.237	47	.49	.25	.112	76.5	Yes
107	98	.02	.65	.23	-.068	143	.63	.23	.137	125.0	No
108	120	.02	.28	.20	.233	48	.27	.20	-.038	78.0	Yes
109	74	.01	.33	.22	.030	119	.31	.21	.043	114.0	No
110	7	.01	.37	.23	.150	79	.36	.23	.201	94.5	Yes
111	33	.01	.48	.25	.174	72	.47	.25	.151	91.5	Yes
112	97	.00	.09	.08	.075	105	.09	.08	-.069	108.5	No
113	148	.00	.31	.21	.121	91	.31	.21	.160	102.0	No
114	106	.00	.27	.20	.109	94	.27	.20	.051	104.0	No
115	48	.00	.28	.20	.080	102	.28	.20	.029	108.5	No
116	125	.00	.17	.14	.048	112	.17	.14	-.012	114.0	No
117	100	-.01	.17	.14	.038	117	.17	.14	.008	117.0	No
118	147	-.01	.13	.11	.232	49	.13	.12	-.043	83.5	No
119	94	-.02	.10	.09	.063	106	.11	.10	-.080	112.5	No
120	42	-.02	.31	.21	.086	101	.32	.22	.043	110.5	No

*$\overline{R}$ is the average rank of $\phi(5,7)$ and $r_{it(7)}$.

APPENDIX 3 (continued)

Rank	Item	$\phi(5,7)$	Diff(7)	$S^2(7)$	$r_{it}(7)$	Rank	Diff(5)	$S^2(5)$	$r_{it}(5)$	$\overline{R}$*	Include
121	89	-.02	.18	.15	-.077	144	.20	.16	.040	132.5	No
122	36	-.02	.56	.25	-.030	136	.58	.24	.076	129.0	No
123	71	-.02	.27	.19	.126	89	.29	.20	.156	106.0	No
124	78	-.03	.41	.24	.005	128	.44	.25	.176	126.0	No
125	26	-.03	.04	.04	-.143	147	.05	.05	-.054	136.5	No
126	76	-.03	.37	.23	.005	129	.40	.24	.022	127.5	No
127	95	-.03	.14	.12	-.180	149	.16	.13	-.043	138.0	No
128	50	-.03	.03	.03	-.041	139	.04	.04	-.156	133.5	No
129	99	-.04	.22	.17	.110	93	.26	.19	-.077	111.0	No
130	60	-.05	.15	.13	.047	113	.18	.15	-.015	121.5	No
131	139	-.05	.29	.21	-.065	142	.34	.22	.024	136.5	No
132	124	-.06	.25	.19	-.121	146	.30	.21	-.006	139.0	No
133	118	-.07	.12	.11	.060	107	.17	.14	.075	120.0	No
134	96	-.07	.14	.12	.014	125	.19	.15	.040	129.5	No
135	142	-.08	.52	.25	.079	103	.60	.24	.225	119.0	No
136	37	-.08	.10	.09	-.052	140	.15	.12	.089	138.0	No
137	67	-.08	.04	.03	-.034	138	.07	.07	-.042	137.5	No
138	133	-.08	.09	.08	.156	77	.15	.12	-.009	107.5	No
139	130	-.08	.36	.23	.123	90	.44	.25	.132	114.5	No
140	65	-.09	.41	.24	.016	123	.49	.25	.100	131.5	No

*$\overline{R}$ is the average rank of $\phi(5,7)$ and $r_{it}(7)$.

APPENDIX 3 (continued)

Rank	Item	$\phi(5,7)$	Diff(7)	$S^2(7)$	$r_{it}(7)$	Rank	Diff(5)	$S^2(5)$	$r_{it}(5)$	$\bar{R}$*	Include
141	30	-.09	.58	.24	.020	121	.67	.22	.081	131.0	No
142	61	-.10	.04	.04	-.156	148	.09	.08	.007	145.0	No
143	38	-.12	.06	.06	.005	127	.13	.11	-.074	135.0	No
144	5	-.12	.24	.18	-.034	137	.35	.23	.073	140.5	No
145	58	-.13	.31	.21	-.009	132	.43	.25	.128	138.5	No
146	111	-.14	.16	.13	.106	98	.28	.20	-.012	122.0	No
147	150	-.15	.14	.12	-.221	150	.26	.19	-.048	148.5	No
148	134	-.15	.10	.09	.022	120	.20	.16	-.038	134.0	No
149	116	-.20	.09	.08	-.057	141	.24	.18	-.009	144.5	No
150	1	-.42	.08	.08	-.003	131	.47	.25	-.065	140.5	No

☆U.S. GOVERNMENT PRINTING OFFICE: 1978-771-122/5

*$\bar{R}$ is the average rank of $\phi(5,7)$ and $r_{it}(7)$.